

76910

From: Rao, Manjunath N.  
Sent: Wednesday, October 02, 2002 9:16 AM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search request for 09/656,640

From: Manjunath N. Rao  
Art Unit 1652, Room 10A11  
Mail Box in Room 10C 01  
Phone: 306-5681

Dat : 10-2-02

Please search the following as soon as possible for application with serial number **09/656,640**

**SEQ ID NO:1, and 3** against all commercial nucleic acid databases including issued patents database and pending application database and provide a **print of all results**.

**SEQ ID NO:2 and 4**, against all commercial protein databases including issued patents database and pending application database and provide a **print of all results**.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.  
Bi technology Patent Examiner  
Art Unit 1652, Room 10A11  
Crystal Mall 1, USPTO.

RECEIVED  
OCT - 2 2002  
CH/CHEN DIVISION  
(STIC)

POINT OF CONTACT:  
PAUL SCHULWITZ  
TECHNICAL INFO. SPECIALIST  
CM1 6B06 TEL. (703) 305-1954

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 10/3  
Date Completed: 10/7  
Searcher Prep/Review: 10  
Clerical: \_\_\_\_\_  
Online time: 10

TYPE OF SEARCH:  
NA Sequences: 2  
AA Sequences: 2  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

✓

**THIS PAGE BLANK (USPTO)**

FORM 1-750-1  
JANUARY 1975  
U.S. DEPARTMENT OF COMMERCE  
BUREAU OF PATENT AND TRADEMARKS









RESULT	3	1958 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AX108672				
DEFINITION	Sequence 1 from Patent WO0121748.				
ACCESSION	AX108672				
VERSION	AX108672.1				
KEYWORDS	GI:13923900				
SOURCE	Stachybotrys chartarum.				
ORGANISM	Stachybotrys chartarum				
REFERENCE	1 (bases 1 to 1958)				
AUTHORS	Convents D.U., Doornink M.U., de Vries C.H. and Wang H.				
TITLE	Detergent compositions comprising phenol oxidizing enzymes				
JOURNAL	Patent: WO 0121748-A 1 29-MAR-2001;				
FEATURES	Location/Qualifiers				
source	1..1958				
	/organism="Stachybotrys chartarum"				
	/db_xref="taxon:74722"				
BASE COUNT	394 a 593 c 490 g 481 t				
ORIGIN					
Query Match	100.0%; Score 1958; DB 6; Length 1958;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 1958;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 ggaatccatcaaatatgatcagccaagctatcggaagccggtggtccttggtgctatc 60				
Db	1 GGATCATCAACATGATGATGACCCAACTATCGGAGCGGTGCTCTGGCTTGTGATC 60				
QY	61 ggcggcagcctcgttcgatgacgacatccgtcttgctgcatcagacagacatgcttcggt 120				
Db	61 GCGCGACAGCTGTGCGATGACGAGATCCGTGCTGATGATGATGATGATGATGATGAT 120				
QY	121 ctacacaagagagcagaagcagctgagctctccctcggccttgatgaagtgctctgcg 180				
Db	121 CTCACCAAGAGGCAACGACGCTGAGCTCTCCCTCCGCTTGTAGGAATGCTTGGCG 180				
QY	181 atccctcctcctbaagggcccaagtagtaagtaacttctatagcttagcagaagccaagt 240				
Db	181 ATCCCTCCTCTBAAGGGCCCAAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 240				
QY	241 tgcataaatctgacgtacgcgtcccaacccaacacacgtggagagagacatctgttactaaga 300				
Db	241 TGCTATATATTCACATACCGTCCCAACCCCAACACTGGAGAGGACATCTGTACTACGA 300				
QY	301 gatgagaattgaagccctctctccacacagatctaacctgatctggagccggccaatagt 360				
Db	301 GATGAGAGATTGAGCCCTTCTCTCCACACAGATCTACCTGTATCTGGAGCCGACCAATAGT 360				
QY	361 tggatagcagatggcaltgcgtcccaagacctaaccatcgtctccctgtyggcaatgaagtgt 420				
Db	361 TGGATAGCATGGCATGTGCTCCCGAGACCTAACATCACTGTTCTGTGGACATGAGAGTGT 420				
QY	421 tgtccgcttcgtgagacagcgagagaaacacccctctcccaagcgtctcaattgcaaggtc 480				
Db	421 TGTCCGCTTCGTGAGACAGCGGAGAAACACCTCTCCCAACACAGCTGCCACTTGGACAGGCTC 480				
QY	481 ttctctcagagctccctcttgatggtctgggcttgagagacactaacccagcctggagataaa 540				
Db	481 TTTCTCTGAGAGCTCCCTTTGATGAGTGGGCTTGAGACACTACCCAGCCTGGGAGATCAAA 540				
QY	541 ggaattactactaaccacaagcagcgtgcccaatgcttggtaaccaatgaacatgcat 600				
Db	541 GGATTACTACTAACCCACAGGACAGGCTGCCGATGCTTGTGATGATGATGATGATGAT 600				
QY	601 gtccataaccgcgcggaagcgtctaatatggttgtaagcctggtgtctatcatgattccagagacc 660				
Db	601 GTCCATACACCGCGGAGAACGCTTACATGAGTCAAGCTGATGATGATGATGATGATGATGAT 660				

OY	661	gqctgagagatgaccccttgaaactcccacagcgcgtacgcgcgagtttgatatacccttgcttc	720
OY	661	ggctgagagatgccccttgaaacctcccacaggcctgacggcgagtttgatatacccttgcttc	720
OY	721	gactgcgaagagatacaacagacagacagcgaactcctctctccacaaatgagagagttccag	780
Db	721	gactgcgaagagatacaacagacagacagcgaactcctctctccacaaatgagagagttccag	780
OY	721	gactgcgaagagatacaacagacagacagcgaactcctctctccacaaatgagagagttccag	780
OY	781	cttcctgaggtgacgttatccaagtgylaaattgagcccatlgagalycttccagtcctag	840
Db	781	cttcctgaggtgacgttatccaagtgylaaattgagcccatlgagalycttccagtcctag	840
OY	841	aagratcgatgatagaatattggtcagatgcttcaacagtgctataccagaagcgttcaagct	900
OY	841	aagratcgatgatagaatattggtcagatgcttcaacagtgctataccagaagcgttcaagct	900
OY	901	tgagccatgagtcacaacgagacagccgcgaagtaacgcttcgcgtctcccaacgctgcgtc	960
OY	901	tgagccatgagtcacaacgagacagccgcgaagtaacgcttcgcgtctcccaacgctgcgtc	960
Db	901	tgagccatgagtcacaacgagacagccgcgaagtaacgcttcgcgtctcccaacgctgcgtc	960
OY	961	tcacgctctctgcgtctgtatccttgtaacctctgagatltcagagaccagacttccctc	1020
OY	961	tcacgctctctgcgtctgtatccttgtaacctctgagatltcagagaccagacttccctc	1020
OY	1021	cagatcatctgcgcgtcgtacgcgtgtgctgtgagagcccggttgacaactgacactctgtac	1080
OY	1021	cagatcatctgcgcgtcgtacgcgtgtgctgtgagagcccggttgacaactgacactctgtac	1080
OY	1081	atccctatgagccgcgcgtcgtgagagtggttatgacttctccacactctgcgtgcagctcc	1140
OY	1081	atccctatgagccgcgcgtcgtgagagtggttatgacttctccacactctgcgtgcagctcc	1140
OY	1141	atcgatataccgacaactcctctggtgctgacaggtctcgtgtgttgagcctgagtttgataac	1200
OY	1141	atcgatataccgacaactcctctggtgctgacaggtctcgtgtgttgagcctgagtttgataac	1200
OY	1201	actgaacagatgacatgagatctcctcctttgataagatcctgagtcgcgcgaacactcttag	1260
OY	1201	actgaacagatgacatgagatctcctcctttgataagatcctgagtcgcgcgaacactcttag	1260
OY	1261	gtgccttgccaactcctcgcgagatgtctccttcccccgcgaagggcgacaacttggaaacccgcgaac	1320
OY	1261	gtgccttgccaactcctcgcgagatgtctccttcccccgcgaagggcgacaacttggaaacccgcgaac	1320
OY	1321	cccaactgatacgagagacttccaccttcgcgcgtgataatgagacagtgagaaatggaaacaaagga	1380
OY	1321	cccaactgatacgagagacttccaccttcgcgcgtgataatgagacagtgagaaatggaaacaaagga	1380
OY	1381	gttaacctctcgcgtgtgcgaagacgctctgcgtccgaatgagtgccccgcgaacactttag	1440
OY	1381	gttaacctctcgcgtgtgcgaagacgctctgcgtccgaatgagtgccccgcgaacactttag	1440
OY	1441	atcgcgcgacttggaaacaactcacaacgcttggagtcacacccctgttcaacttaacctgct	1500
OY	1441	atcgcgcgacttggaaacaactcacaacgcttggagtcacacccctgttcaacttaacctgct	1500
OY	1501	gaactcgcagatccttctcgttccaaatgccgtgtgagtcgagaccttaatgagagctgcgtgct	1560
OY	1501	gaactcgcagatccttctcgttccaaatgccgtgtgagtcgagaccttaatgagagctgcgtgct	1560
OY	1561	ctcaagagatgtgtgcgtgcgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt	1620
OY	1561	ctcaagagatgtgtgcgtgcgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt	1620
OY	1621	ccttcgcgaagatctcgcgccttccaaacgaggttttccactatagtcaactaactaaca	1680
OY	1621	ccttcgcgaagatctcgcgccttccaaacgaggttttccactatagtcaactaactaaca	1680
OY	1681	gtggtgtcatactgttccactgtcacaacactgtatccacaagagacacagacatgagtgctg	1740
OY	1681	gtggtgtcatactgttccactgtcacaacactgtatccacaagagacacagacatgagtgctg	1740
OY	1741	cttcaatgtaoactgltcgtgltgactatggtcacaactataccgagatcalttagccca	1800



Db	1627	CAAGATGTTGTCGTGGTGGCTGCTGTCGTGAGTTGTCATGTTTGAGGCCCACTACGCTCC	1686
Qy	1623	tttcccgtaagttccgcgctttaccctaactggttttcaactcaactcaactcaagt	1682
Db	1687	TTTTCCCGTAAGTTCTCGGCTTTTACCTAACTGTTTTCATCTATCTCTAACATCTCAAGT	1746
Qy	1683	ggtgtcctacatgttgcactgcccacaacccgatactcaacgagagaccagacatgagtctgt	1742
Db	1747	GGTGCTCTAATTTTGACATCGCCACCAACCTGATCCACGAGACCACGACATGATGGCTGCT	1806
Qy	1743	ttcaatgtcaactgtttccggtgactatggtcctaacaactacaaccgagttcaattgaccacatg	1802
Db	1807	TTCAATGTTCACAGTCTCTGGTGACTATGGCTCAACTACACCGAGTTTCATTGACCCCATG	1866
Qy	1803	gagcgtctcttggaagccccgcgcctctctccctcggaaggtttcgaaatgaatggtccggtgac	1862
Db	1867	GAGCCTCTCTGGAGGCCCGCCCTTCTCCCTCGGAGAGTTTGAAATAAGGCTCGGGGTAC	1926
Qy	1863	ttcagcgaagcttgcacatactgaacgcgattcaagaagatggtcctagcttcaacccctacgc	1922
Db	1927	TTTCACGAGAGCTTGCATCACTACACCGCAATTTCAGAGATGGCTTCAACCCCTACGCC	1986
Qy	1923	cagcctgatgatgatgatacgccgctggaagataga	1953
Db	1987	CAGGCTGATGATGATGATCCGCTGAGAGATAAA	2017

RESULT	5								
LOCUS	AX101098								
DEFINITION	AX101098								
ACCESSION	Sequence 3 from Patent WO0121809.								
VERSION	AX101098								
KEYWORDS	AX101098.1								
SOURCE									
ORGANISM	Stachybotrys chartarum.								
REFERENCE	Stachybotrys chartarum.								
AUTHORS	Eukaryota, Fungi; Ascomycota; mitosporeic Ascomycota; Stachybotrys.								
TITLE	1 (bases 1 to 2095)								
JOURNAL	Wang, H.								
FEATURES	Stachybotrys phenol oxidizing enzyme								
source	Patent: WO 0121809-A 3 29-MAR-2001;								
	INTERNATIONAL, INC. (US)								
	Location/Qualifiers								
	1..2095								
	/organism="Stachybotrys chartarum"								
	/db_xref="taxon:74722"								
BASE COUNT	437 a								
ORIGIN	618 c								
	510 g								
	530 t								

	Query Match	Best Local Similarity	99.4% 99.8% Matches 1948;	Score 1946.2;	DB 6;	Length 2095;
			Conservative	0;	Mismatches 3;	Indels 0;
					Gaps 0;	
QY	3	atccatcacatgatcacaagaatcatctcggagccgtggtctctgggacctgtgcatcg	62			
Db	67	AACGATCAACATGATACCAAGCATGCGAGCGCTGCTGGGCTTGGCTGTGATGG	126			
QY	63	cggcagctctgtcgatgcagatccgttgcgtgtgatcatgacaagaatgcttcogtct	122			
Db	127	CGGAGCCTCTGCATGCGACATCCGTTGCTGTGATGACAGACATGCCCTCCGGCT	186			
QY	123	caccagaagagacagcgcagctgtagtccctcccttgaccttgaagaagtccttcgcgat	182			
Db	187	CACCAAGAGCGACAGCGCAGCTGACTCTCCCTGGCCCTTGTACGAAGGCTCTGGCCGAT	246			
QY	183	ccctctcttgaaggcgcgcccaagtagtaagtaacattcctaaggtctagcagagccaaagtctg	242			
Db	247	CCCTCTCTTGAAGGGCCCAAGTAGTAGTCAATTCTATTAGGTACGAGACCAACGTTG	306			
QY	243	ctaatcatctgaagtaacgttccccaaccccaaacctgtgagagacaatcttgtactacaga	302			

Db	307	CTAATCATTTGCAGTAACTGCTTCCCAACCCCAAACTGGAGAGACATCTTTACTACGAGA	366
Qy	303	tggagataaggccctctccaccagatactaccctgaatctggaagcccgccaacatggttg	362
Db	367	TGGAGATTAAAGCCCTTCTCCACCAAGATCAACCTGTATCTGGAGCCGGCAACATGATTGG	426
Qy	363	gataagatgagatgtccccaagacctacataatgtctccctggtgacctggagatgttg	422
Db	427	GATACGATGGCATGTGCTCCCAAGACCTTACATCATCTGTTCTGTGGCACTGAGAGTGTGG	486
Qy	423	tcgcgttcgtgaacagcggagagaaacacctctcccaacacgcttcacatctgcaagtcctt	482
Db	487	TCCGCTTGTGTAAACAGCGGAGAGAACACTCTCCAAACAGCGTCCACTTGACAGCGCTTT	546
Qy	483	tctctcgagctcccttttgaigtgttgagcttgaagaacctcccaagccttggcagatgaag	542
Db	547	TCTCTCGAGCTCCCTTTGATGTGTGGCTAGGACACTACCCAGCCTGGCGAGTCAAAAG	606
Qy	543	atatactaccccaacaagagaagctgccgaatgtcttgttgatacctatgacatactgcaat	602
Db	607	ATTACTACTACCCCAACAGGAGGCTGCCGATGCTTTGATACATGACATGCGCATGT	666
Qy	603	ccatcacccgcccgaagaagcctataatgggtcgaagctgtgtgtctacatgataccagaaccg	662
Db	667	CCATACCCGCCGAAGAACGCTTACATGGGTGAGGCTGATGCTTACATGATCCAGACCCGG	726
Qy	663	ctgaagatgcccctgaacctcccccagcggctcagaagcagattgaatacccccttggtctga	722
Db	727	CTGAGAGATGCCCTTAACCTCCCAAGCGGCTACGCGGAGTTTGAATATCCCTTGTTCTGA	786
Qy	723	ctgccaaagcgtatacacgacgacgcactctctctccaccaatlgagaagattccagct	782
Db	787	CTGCCAACCGATTACAGCGACAGCGGCACTCTCTTCTTCCACAAATGAGAGAGTTTCCACT	846
Qy	783	tctggggagagctattccaagtgtagtgaagttgaagcccatlgaaatgcttaagatccataaa	842
Db	847	TCTGGGGTGAAGCTTAATCAAGTGAAGTTAGACCCATTTGAGATGCTTACGATCTCAGAA	906
Qy	843	glatcgaatglatgaatltgtgcatctgctctaacacagtgcatctacagaagaacgctcaagct	902
Db	907	GTATGATGATTAAGAAATTTGATGATGCTTACAGTCAACAGATGATACAGAACGTCACGCTTG	966
Qy	903	gacctatgctcaactgtcagcgcgcgcgcgaagtacacgcttcgcctctccaaagctgcgcgtctc	962
Db	967	GCTTATGCTCAACCTGAGCGCGGCAAGTACCGCTTCCGCTTCTCTCAACGATGCGGCTGTC	1026
Qy	963	acgcctcttcgctctgatactctgcaacctctgaggaatcaagaagacaagctccctctca	1022
Db	1027	ACGCTCTTGGCTCTGTATCTTCTGTAACCTCTTGAGATTCAGAGACCAAGACTTCCCTTCA	1086
Qy	1023	ggtcaattgcgcgtgaacggtgtgtctgtcttgaaggccctgttgaacatgacatctgttacct	1082
Db	1087	GGTCAATTGCGCTGACCGTGTGCTGTTGAGGGCCCTGTTGACACTGACACTGACTGTATCAT	1146
Qy	1083	ctctaatggccagagctctgggaagtttlttaacagctctccacctctgcgtcggcgcaatgcat	1142
Db	1147	CTCTAATGGCGAGGCGCTGGGAGGTTGTTATCAGACTTCTCACCTTGCGTGCCCACTCAT	1206
Qy	1143	cgatatccgaacaccttcctgtgtctgaagctctcgtgtgttgaacatgagttgataaacac	1202
Db	1207	CGAATATCCGCAACCTTCTGCTGCTGACGAGTCCGCTGTTGAGCCGAGATTGATTAACAC	1266
Qy	1203	tgaagaagctatgagatctgttgaatgaagtcccttgaatcttgaccccggaaccttctgaagt	1262
Db	1267	TGACAAAGGTATGAGATTGCTGCTGTATGAAGCTCTTGAAGTGGCCCGCACTTCTGAAGT	1326
Qy	1263	gctctgcaaacctccgaatgttctcttcccccgaagtcggaacatlgggaccgcgcaaaccc	1322
Db	1327	GCTGCGCAACCTCGAATATTTCTTTTCCCGAGGGCGGCAACTGGGAACCCCGCAAAACC	1386
Qy	1323	caatgatagcagagacttcaactcttgccgtgcgtgtctaaatlgagacatgacaatcaacgagt	1382
Db	1387	CACGATGAGAGAGACTTTTCACTTCTCGCGCGTGTATATGACAGTGAACATCAACGAGAGT	1446

Page 6	
366	A
362	J
426	S
422	J
486	S
482	L
546	T
542	J
606	S
602	L
666	T
662	J
726	S
722	A
786	A
782	L
846	T
842	A
906	A
902	J
966	S
962	S
1026	S
1022	A
1086	A
1082	L
1146	
1142	L
1206	
1202	S
1266	S
1262	L
1326	
1322	S
1386	L
1382	L
1446	

[illegible]

Db	127	CGGACGCTCTTCATGCGAATCCGTTGCTGGTCCATGCAAGACATCTTCGGGTCT	186
Qy	123	caaccaagagcagacgcagcgtgagtcctccctcgtgccttgttacaagtgctcttgcgat	182
Db	187	CACCAAGAGCGACAGCAGCTGAGTCTCTCCCGCTTGATGAGAGTGCCTTCGCAT	246
Qy	183	ccctccctctgaagcgccccaagttagtaagtaactctctatagctctgcaagccaacgttg	242
Db	247	CCCTCTCTTGAAAGCGCCCAAGTAGTAGACATTCTTAAGGCTACCAAGCAACGTTG	306
Qy	243	ctaactcattgcaagcagctccccaaccaaccaactgagaggaacattgttactacagga	302
Db	307	CTAAATCATTTGCAAGTACCGTCCCAACCCCAACACTGGAAGAGACATCTTGTAACGAGA	366
Qy	303	tggagatgaagccctctctccacaagaactacactgtatcttgagccgagcaaatggttg	362
Db	367	TGGAGATTAGGCCCTTCTCCACCAAGATCTACCCCTGATCTGGAGCGCGCAACATGTTG	426
Qy	363	gataagatgagcatgtccccaaggaactcaactatgtctctcgttgcaatgagtggtg	422
Db	427	GATAGAGTGGGATGTCCCCAAAGACTTACATCTGTTCTCTGTGGCACTGGAGATGTTG	486
Qy	423	tccgctcttgaaagacgaggaagaacaacctctccaagagctccacttggaacgcctctt	482
Db	487	TCCGCTTGTAACAGCGGAGAGACACTCTCCAAACAGCGTCACTTGCAAGGCTCTT	546
Qy	483	tctctcgaagccctcttgaatggttggtgagtagaacaactccaagcccttgcagatcaag	542
Db	547	TCTCTCGAGCTCCCTTGATGTTGGGCTTAGGACACTACCCAGCCTGGCGAGTAGTAAG	606
Qy	543	attactactaccaacaagcagcgtgcgcgcgaactgcttgtgtacaatgacatgcgaatgt	602
Db	607	ATTACTACTACCCCAACAGGAGGAGCTGCCCGACTGCTTGTAGTACATGACATGCCATGT	666
Qy	603	ccatcacgcgcggaagacgcctacatcttggtaggctggtgtctacatgataccagaccgg	662
Db	667	CCATCACGCGCGGAAGACGCTCACTAGGAGTGAAGCTGTGTCTACATGATCAAGACCCGG	726
Qy	663	ctgaggaatgcctctgaactccccaagcgcgtacagcgagttgatatccctctgtcttga	722
Db	727	CTGAGAGTGGCTGTAACTCCCAAGGGGTAGGGCAGTTGATATCCCTTGTTGTGA	786
Qy	723	ctgccaagagatatacagcgagaagcaactctctctccaccaatgtagaggttccagct	782
Db	787	CTGCCAAGCGATTACAAGCGACGCGCACTCTTCTCCACCAATGAGAGAGCTTCCACCT	846
Qy	783	tctggggtgaaggtattcatcaagtgtgtaagttgagcccaattgagatgcttcaatcctaga	842
Db	847	TCTGGGAGGAGTTAATCAAGTGTAAAGTTGAAGCCCATTAAGATGCTTCAAGTACCTAGAA	906
Qy	843	gtatcagatgtaagaatttgcataagtcactcaacagtgctataccaagaagctcaaccttg	902
Db	907	GTAATGATGTATGAATTTGTGCATGCTCTTAACAGTGTATCAACMAAGGTCACACTTG	966
Qy	903	gactatgctcaacgttgacgcgcgccaagtacacgcttcgccttccctaagcgtgcgctctc	962
Db	967	GCGTATGCTCAACGTGCACCGCGCGCAAGTACGCTTCCGCTTCTTAACGCTGGCGTCTC	1026
Qy	963	acgcctcttgctctgtatcttgcttactcctltgagaatcaagaaaccaagacttccctcca	1022
Db	1027	ACGCGTTTCCCTCTGTATCTTGTACTCTTGAGGATTCAGAGACCAGACTTCCCTTCCA	1086
Qy	1023	ggtcaatgcgcttgaaggtgtctcgtcttgagggcccttgtgaacacgnaacactttaat	1082
Db	1087	GGTCAATTGCCCTACAGGTGGTGTGCTTGAAGGCCCTGTGACACTGACACTCTTATAT	1146
Qy	1083	ctctatgcccagagcctgagaggtgtgttlaacgaacttccacacttgcgtgcgaacttcac	1142
Db	1147	CTCTATGGCCGACGGCGTGGGAGGTTGTTATCAGACTTCTCCACTTTCGCTGGGCATCCAT	1206
Qy	1143	cgataatccgaaacttccctggtgctgaaggtctcagtlgtgagcctgagttgtataaac	1202

Db	1207	CGATATTCGGCAACCTTCCTCGTGGTGTGAGCGTGTGAGTGTGAGCTTGAGTTGATTAACAC	1266
Oy	1203	tgacaagatcatcgcaatttcgctcgtgtgtagaagcttccttaagtcgcccagacattcgaagt	1262
Db	1267	TGACAAAGGTCAATGCGCATTCCTGCTGTGTGATGAAATCTCTTGAGTGGCCCGACACTTCCTGAGGT	1326
Oy	1263	gctctcaaacctccgagatgtctccttcocccgaggcgccaaactctggagcccccgaacc	1322
Db	1327	GGCTCCCAACCTCGGAGATGTTCTTCTTCCCGAGGGCCGCAACTGGGACCCCGCAACC	1386
Oy	1323	cactgaatgacgaagacttcaaccttcgcccgtgtctaaatgacagatgacaaatcaacgaagt	1382
Db	1387	CACATGATGACGAGACTTTCACCTTGCGCCGCTCTAATGAGACAGTGGACATCAACGAGGT	1446
Oy	1383	tacctcttcgagatgacgagaaacgcgtctctccgagatgtcccgcggaacagtgtgaat	1442
Db	1447	TACCTCTTGCGGATGTCGAGAACCGTCTCTCCGCAATGTCGCCCGCACTGTTGAGAT	1506
Oy	1443	ctgaggaacttgagaacaactccaaaggttgagactaacctgttcaacatcacctgttga	1502
Db	1507	CTGGCGACTGTGAGAACAACTCCAACGGTTGGACTCACCTGTTCAATTCACTTGTTGA	1566
Oy	1503	cttcggaatccttcttcgcttcacatgcccgtgagatcgaaacctatgagctgtgtctc	1562
Db	1567	CTTCCGAGCTCTTCTTCTGCTCCACAGCCCGCTGAGATCGACCTTATGAGGCTGAGCT	1626
Oy	1563	caagaatctgtctgagcctgagcctgcgtcgtagagtgtctatgttgagagccaaactacgtcc	1622
Db	1627	CAAGATGTTGTGCTGGCTGGCTGCTGCTGAGAGTGTGTCTATGTTGAGGCCCACTAGCTTC	1686
Oy	1623	ttcccgtaagtcttcgccttctacctaactggttlltaactcaatgactaacatacaagt	1682
Db	1687	TTTCCCGTAAGTTCTCGGCTTTTACCTTAACGTGGTTTAACTCATCTTAACATCTCAAGT	1746
Oy	1683	ggtgtctaacatgttgacatgcccacaacctgatacagaaggaccagacatagtatgctact	1742
Db	1747	GGTGTCTACATGTTTGCACTGCCACAACCTGATCCACAGAGACCAGACATATATGCTGCT	1806
Oy	1743	ttcaatgtcaactgtctcgtgtgataatgagctaacatacaaccgaattcaatgaccatg	1802
Db	1807	TTCAATGTACACTGTTCTCGGTACTATGAGTCACTACACCACTGATTCATTGACCCCATG	1866
Oy	1803	gagcctctcttgaggccccgcgcctctcctccctcgagaaggttgaaatgagctcggtgtac	1862
Db	1867	GAGCTCTCTGGAGGCCCGCCCTTCTCTCTCGGAGAGTTTGAAATGAGCTCGGGTAC	1926
Oy	1863	ttcagcagagcttgcatctactactgacccgatactaaagagatgagctagtttcaacccctacgc	1922
Db	1927	TTCAACGAGCTTGCCATCATCATGACCGCAFTGCGAGATGGCTAGCTTCACCCCTACGCC	1986
Oy	1923	cagagctgatatgatatgcccgcctgaggaagtaga	1953
Db	1987	CAGGCTGATGATGATGATGCCGCTGAGAGTAA	2017
RESULT	7		
MYROR	LOCUS		
DEFINITION	MYROR	1942 bp	MRNA
ACCESSION	M. verrucaria mRNA for bilirubin oxidase, complete cds.		linear
VERSION	D12579.1	GI:456711	
KEYWORDS	bilirubin oxidase.		
SOURCE	Myrothecium verrucaria (strain:W-1) cDNA to mRNA.		
ORGANISM	Myrothecium verrucaria		
REFERENCE	Eukaryota, Fungi; Ascomycota; mitosporic Ascomycota; Myrothecium.		
AUTHORS	1 (bases 1 to 1942)		
TITLE	Kojikeda, S., Ando, K., Kaji, H., Inoue, T., Munao, S., Takeuchi, K. and Samejima, T.		
JOURNAL	Molecular cloning of the gene for bilirubin oxidase from		
MEDLINE	Myrothecium verrucaria and its expression in yeast		
REFERENCE	J. Biol. Chem. 268 (25), 18801-18809 (1993)		
	2 (bases 1 to 1942)		

**AUTHORS** Ando, K.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (07-JUL-1992) Keiichi Ando, Amano Pharmaceutical Co., Ltd., Tsukuba Research Laboratories; 22 Miyukiyaoka, Ibaragi 305, Japan (Tel.:0298-56-5026, Fax:0298-56-5012)  
On Feb 26, 1994 this sequence version replaced gi:436235.

**COMMENT** Submitted (07-JUL-1992) to DDBJ by:  
Keiichi Ando  
Tsukuba Research Laboratories  
Amano Pharmaceutical Co., Ltd.  
22 Miyukiyaoka  
Ibaragi 305  
Japan  
Phone: 0298-56-5026  
Fax: 0298-56-5012.

---

<b>FEATURES</b>	Location/Qualifiers
source	1..1942
	/organism="Myrothecium verrucaria"
	/strain="MT-1"
	/db_xref="taxon:5532"
sig_peptide	66..179
CDS	66..1784
	/EC_number="1.3.3.5"
	/codon_start=1
	/product="bilirubin oxidase"
	/protein_id="AAA02123.1"
	/db_xref="GI:456712"
	/translation="MEKRTTCAALSLLENSNAOVAQSPRETPATCHLFRVAOISP OTPEPTVPIPPYKOPRLTVNPVNQEELWYELEIKPPTHROYIPDLGSADLVGIDG MSDEPTFOVPRGVEIVRFINNAEAPSVHLHSGFSAAFDGAEDITLEGSRKYD KNRSARTLWYHDHAMHTAEANA YRQAQSGLMLTDPAEDLNLP SGGEFDDIMILYS POYAAGNLTVMTEENLSFWGDYHVHGQGPMPKRNPRRYREFFLDAASRFGLYF ADEPDADRPLPFVKVIASDSLFLEHPADSLLTYISMAERYEVDFSDYAKTIELNL GGSTIGTGDTDYDNMTRKVRFPVAVDPTQDPDRSYVANLRDVPSPTNTPEPROFE GRCPETATVINGVAFAFDONRIIANPVGYERRELINAGCMTHPHIHLDKVISR TSGNALTATVPYESGLKDQVMYLGRREVYVBAHYAPPQCYTMCHNLIEDDMAAF FNATVELPDGYINATVFVDPMEEIQARPYDELGFQAOOGFVOAVTERIQIAEYRP YAADP"
mat_peptide	180..1781
	/product="bilirubin oxidase"
polya_signal	1832..1837
polya_site	1942
BASE COUNT	440 a 549 c 436 g 517 t
ORIGIN	

```

Query Match      27.0% ; Score 528.6; DB 8; Length 1942;
Best Local Similarity   62.0% ; Pred. No. 1.5e-114;
Matches 1058; Conservative    0; Mismatches 474; Indels 175; Gaps       7;
```

---

Oy	239	gttgctaatcatltygaatccgcttcccccaaccacaacattgagagaagatatctglactaac	298
Db	237	GTTAGCAGCCCGCTGTACTGTAAACCAATCTGTGAATGACAAGAAGATTCTGTACTAT	296
Oy	229	gagaigagagttagggcctttcccaccaaggattcaccttgttcggcgccggccaacatg	358
Db	287	GAGGTGCAGATCAACCCTTACTCATCACAGGTTTTAACCTTGACCTTGGATCCGCTGATCTG	356
Oy	359	gttggatacgaatygcatalytccccaggaacctaacltaactcgttccctgtgagcayagat	418
Db	357	GTCGGGTATGATGGAATGTCCTCGGCCCTTACTTCCAGGTCTCTCGTAGGTGAANACA	416
Oy	419	gtttctcgctctcygaacagcgysagagaacaacctctcccaacagcgfctccactlycaogc	478
Db	417	GTTTGCCCTTCATTAAACAATGCTGAG-----GCTCCTTAATCCGTTCACCTCCACGGA	470
Oy	479	tcttctctcgagctccctctgatgattgaggctlgagcaactaccagccctgcgagtac	538
Db	471	TCATTCCTCTGCTGCCGCTTTGACGGATGGGCAAGAGACATACACGACCTCGGCAGCTTC	530
Oy	539	aagattactactaccaccaagcgagctgcccgcgatcttgggtacacatygacacatg	598





[illegible]

Db	1330	TGAGACGTGGGAGCGTCATCAACGGCCGGTAACGCTTGGAACGCACGCCATTTCACATCCATC	1389
Oy	1496	tcgltgactccagblccttcgtlctc-----cactgcgccgyagatcgagcctt	1546
Db	1390	TGTGTGACTTCAAGGTGATTTCTCTACTTCCGGCACAAACGCCGCGACAGTCATGCCAT	1449
Oy	1547	atgaagcgtcgtgtctcaagaagtgtgtctgacgtgcgtcgttgtagtgltgtclatgtgt	1606
Db	1450	ACGAG---TCCCGTCTCAAAGACGTTGCTGGCTTGCTGCGCTGAACCTGtGTTGTTG	1506
Oy	1607	aggcccaactacgtcctttcccgaagtlctgccttlaactaacigtlttcaactcat	1666
Db	1507	AAGCATTAATACCAGCGCTTTCCC-----	1528
Oy	1667	gctaacaactacaagtggtgtctacatcgtltgcactgcacaaacctgatccaagagacca	1726
Db	1529	-----TGGGTATACATGTTCCATTGCCCCAATAATTGATTCAAGSAGATCA	1573
Oy	1727	ggacaatgtagctgtcttcaaatcgttaactcgtctcgtgtgactatggtctcaactaaccca	1786
Db	1574	CGATTATGATGCTGCTTTAACGCCACCGCTCTGCGCAATTATGGCTTAAATGCCACTGT	1633
Oy	1787	gttcalttaccccaatgatgagcctctctcttggaagccccccctcctcctcctcggagatgtca	1846
Db	1634	TTCGTGTGACCCCTTAGGAAGACCTTTGGCAGGCGTCGTCCTTATGAACTCGGGGAATTCCA	1693
Oy	1847	gaatgctcgtcgggtgacttcaagcagagctgtgccatacgtacagcatcagaagaagtctag	1906
Db	1694	GCCCTAGAATGGCCAGTTCAGCGCTTCAGGCGCTTACTAGCGTATTTCAGACTATGCTGTA	1753
Oy	1907	cftcaaccctacagcccgagctgatga	1933
Db	1754	ATACAGACCTTAGCGCCGACGCTGACGA	1780

RESULT 9  
AMU271104

LOCUS	AMU271104	2126 bp	mRNA	linear	PLN 08-JUN-2001
DEFINITION	Acromonium murorum mRNA for polyphenol oxidase (ppoa gene).				
ACCESSION	AJ271104				
VERSION	AJ271104..1	GI:6996277			
KEYWORDS	oxidase; ppoa gene.				
SOURCE	Acromonium murorum.				
ORGANISM	Acromonium murorum				
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Acromonium.				
AUTHORS	1 (bases 1 to 2126)				
TITLE	Gouka,R.J., van Der Heiden,M., Swarthoff,T. and Verrips,C.T.				
CITATION	Cloning of a Phenol Oxidase Gene from Acromonium murorum and Its Expression in Aspergillus awamori				
JOURNAL	Appl. Environ. Microbiol. 67 (6), 2610-2616 (2001)				
MEDLINE	21268855				
REFERENCE	2 (bases 1 to 2126)				
AUTHORS	Gouka,R.J.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-JAN-2000) Gouka R.J., Biotechnology, Unilever Research Vlaardingen, Olivier van Noortlaan 120, 3133 AT Vlaardingen, NETHERLANDS				
FEATURES	Location/Qualifiers				
source	1..2126				
	/organism="Acromonium murorum"				
	/strain="CBS 157.72"				
	/sub-species="murorum"				
	/db_xref="taxon:45278"				
	15..97				
5'UTR	/gene="ppoa"				
gene	15..2120				
mrna	/gene="ppoa"				
	15..2120				
	/gene="ppoa"				
	/product="polyphenol oxidase"				
	98..160				
	/gene="ppoa"				

sig\_peptide



```
CDS
98..1906
/gene="ppoa"
/codon_start=1
/product="polyphenol oxidase"
/protein_id="CAB75422.1"
/db_xref="GI:696278"
/db_xref="SPTREMBL:Q9P8C3"
/translation="MVATPALRALALVLSLKAQAMKPEFLDIPPEEAAALAIAYED
PANDORSPANDLORSPALPAYLFOAPISIPVKEPLFVTNPNYNGEIDYEIE
IKHSEQVFPDLPADLYGDISPSPFIDKEGESVRYNKAIVTSSTLHHSFS
RAPDMGAEVNTNEEYKDYYPNAOARFSWYHDHAEHETAEKNVNCAGYTIIDP
AEDSLGPSYGEYDPLILTSKOYNSDGLSTKQTSQIMGDVIOVNPWPVEDY
EPKRYRFLDLAAVRSFSLYFVDPADEDTRIPEQVIAVSDSLPEPVTSLVISA
ERYEIIDFDESEKTEILRNEPAVGGLIEVNDTDKVMFNVAEGPLSPDYSV
PSTLRDVPSPSTSTTIDHSFRFARTAGOWSINCVTFSDVBRILLANVPIGVOLMOT
TMAAGRTHPHIHIVDKILSRCAATGATRTREYEPESAGLRKVYLKRETYVE
AHYAPWGVYMFHCHNLTHEDDMATNVLVPGYNSTALADPMODEFRAPKYVD
NDVEVERENAFSTDEIEAOVDMASYLPYDNDPMEIV"
161..280
/gene="ppoa"
/feature
misc_feature
161..280
/note="polyphenol oxidase propeptide"
mat_peptide
281..1903
/gene="ppoa"
/evidence="experimental"
/product="polyphenol oxidase"
3'UTR
1907..2120
/gene="ppoa"
2073
polyA_site
2073 c
BASE COUNT 442 a 723 c 546 g 415 t
ORIGIN
```

```
Query Match 26.4% Score 517.4 DB 8: Length 2126:
Best Local Similarity 61.4% Pred. No. 6.8e-112:
Matches 1041: Conservative 0: Mismatches 476: Indels 178: Gaps: 6:
98..1906
/gene="ppoa"
/codon_start=1
/product="polyphenol oxidase"
/protein_id="CAB75422.1"
/db_xref="GI:696278"
/db_xref="SPTREMBL:Q9P8C3"
/translation="MVATPALRALALVLSLKAQAMKPEFLDIPPEEAAALAIAYED
PANDORSPANDLORSPALPAYLFOAPISIPVKEPLFVTNPNYNGEIDYEIE
IKHSEQVFPDLPADLYGDISPSPFIDKEGESVRYNKAIVTSSTLHHSFS
RAPDMGAEVNTNEEYKDYYPNAOARFSWYHDHAEHETAEKNVNCAGYTIIDP
AEDSLGPSYGEYDPLILTSKOYNSDGLSTKQTSQIMGDVIOVNPWPVEDY
EPKRYRFLDLAAVRSFSLYFVDPADEDTRIPEQVIAVSDSLPEPVTSLVISA
ERYEIIDFDESEKTEILRNEPAVGGLIEVNDTDKVMFNVAEGPLSPDYSV
PSTLRDVPSPSTSTTIDHSFRFARTAGOWSINCVTFSDVBRILLANVPIGVOLMOT
TMAAGRTHPHIHIVDKILSRCAATGATRTREYEPESAGLRKVYLKRETYVE
AHYAPWGVYMFHCHNLTHEDDMATNVLVPGYNSTALADPMODEFRAPKYVD
NDVEVERENAFSTDEIEAOVDMASYLPYDNDPMEIV"
161..280
/gene="ppoa"
/feature
misc_feature
161..280
/note="polyphenol oxidase propeptide"
mat_peptide
281..1903
/gene="ppoa"
/evidence="experimental"
/product="polyphenol oxidase"
3'UTR
1907..2120
/gene="ppoa"
2073
polyA_site
2073 c
BASE COUNT 442 a 723 c 546 g 415 t
ORIGIN
```







DEFINITION	Streptomyces aureofaciens type-II polyketide gene cluster, partial sequence.	gene
ACCESSION	AY033994	CDS
VERSION	AY033994.1	
KEYWORDS	GI:14486270	
SOURCE	Streptomyces aureofaciens.	
ORGANISM	Streptomyces aureofaciens	
REFERENCE	Bacteria: Firmicutes; Actinobacteria: Actinobacteridae; Actinomycetales; Streptomycetales; Streptomyces. 1 (bases 1 to 9531)	
AUTHORS	Kormaneec, J., Bistakova, J., Novakova, R., Homeroova, D. and Rezuchova, B.	
TITLE	Cloning and characterization of a new polyketide gene cluster in Streptomyces aureofaciens CCM3239	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 9531)	
AUTHORS	Kormaneec, J.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-MAY-2001) Gene Expression, Institute of Molecular Biology, Slovak Academy of Sciences, Dubravská 21, Bratislava 84251, Slovak Republic	
FEATURES	Location/Qualifiers	
source	1. 9531	
	/organism="Streptomyces aureofaciens"	
	/strain="CCM3239"	
	/db_xref="taxon:1894"	
gene	<1. 212	
	/gene="aur2j"	
CDS	<1. 212	
	/note="aur2j"	
	/codon_start=3	
	/transl_table=11	
	/product="aminotransferase-like protein"	
	/protein_id="AAK61712.1"	
	/db_xref="GI:14486271"	
	/translation="IYTFRYAPLHKVPTGTPTDRLPSSDWSDRTLCLPLHGLSDADVLTVAAASLRKAVEATAEKNARS"	
	complement(345. 1772)	
gene	/gene="aur2i"	
	complement(345. 1772)	
CDS	/gene="aur2i"	
	/note="Aur2i: similar to Bacillus subtilis spore-coat protein Cota"	
	/codon_start=1	
	/transl_table=11	
	/product="oxidoreductase-like protein"	
	/protein_id="AAK61713.1"	
	/db_xref="GI:14486272"	
	/translation="MAAAGAAALFPMDRITGASQAVASSLATAEAPITVPEAHAMP LPKNIKPTSTFATSDLYEIRMOEAQVEIVKIGMSKVITYGTGPGPIIRATOGREVVV ROINELQVNTAVHLGHAHVSEHDGLPMQITIVRGERTYRPNQPAASLMDHHAH LEADNFMGLHGLILTDHNERLPLPSGPDVPLVIRDAVEADGTLITRSDCPH MLVNGKERPFQVAAKRYRFRVYNACANRYKLFADGIEFTQIGTGGLFEPVQS ELMALGERADIVDESRYKVDVSVLENQAOSIERPEVAREPDIIVTADYSFVR LTYTPQPTVEDEFEIRTFEPAMTINGOSYDPMRVDTAKLGTTEVATRVNVAAPA BGDDEFLHSHFHLTYERVLERNKPACTROLDGLKDYTLGPGDIVKIAMTMGPPT GQYLHCHQLHSSSGOMGRIDYA"	
	complement(1782. 2513)	
gene	/gene="aur2h"	
	complement(1782. 2513)	
CDS	/gene="aur2h"	
	/note="Aur2h"	
	/codon_start=1	
	/transl_table=11	
	/product="unknown"	
	/protein_id="AAK61714.1"	
	/db_xref="GI:14486273"	
	/translation="MSVAREIVAAAGRRAGTSGNGPCVAGRGREDDPGGAPCA VPGVAGRGAGRGALHLISLSVPTGLLAFQAFVANIYTRTEAPPPQEDNHW WHSGVPEPEAFRPGFLATVAFTVLTPLLAGOVLRRRCGLVAGAGRSIPADL GFAAAATVAGAVTVPLVLAFAFSVDGSGPLGAVTVADAVTAMRLCALALLAFVA GAWMAAGAPSLRREGSLHAESS"	
	complement(2510. 3367)	
gene	/gene="aur2g"	
	complement(2510. 3367)	
CDS	/gene="aur2g"	
	/note="Aur2g"	
	/codon_start=1	
	/transl_table=11	
	/product="oxygenase-like protein"	
	/protein_id="AAK61715.1"	
	/db_xref="GI:14486274"	
	/translation="MSDRTERWTLTASRGCLGALAEVYLEAGDTVYATARSIPALD GLVERFGEERVPALDVTIDRAAVLAIAVAEAERTGRIDVLLNNAGYGLAAVEEVNT QVRDQDFVNFEGVALMTVOAVLPVRRQSGSHLQVSSIACTYTPNLGLCAKSWALE AVSETTIAOEAEGFVKTVLVEAGEFRTDMGASGVVATPFAVDEVLAKRHGLSGAY AHLQDPDRKAGEALTLKTDPAEPPLVLLGEGAADLAPVTLARLARGWEEMDLART TDPARESGP"	
	complement(4349. 5134)	
gene	/gene="aur2e"	
	complement(4349. 5134)	
CDS	/gene="aur2e"	
	/note="Aur2e"	
	/codon_start=1	
	/transl_table=11	
	/product="cyclohydrolase-like protein"	
	/protein_id="AAK61716.1"	
	/db_xref="GI:14486275"	
	/translation="MTQGLREVEHEITVSAPAAVYRLIAEVVNMPIPPITYVDO VPBEGGERIRIATANGAEKKNSTRTLDALRIETPQEVTPAPPAANGTIIIE LSADSVRLDHDTRAVDDSDSLAWIDEADVRSSELPAKLVNDAHAEBELITS FEDVQJEGSKADVDYDFVNEADRWHERLPHVATVRLREDDPGLQELMDVRAKGSVH TTKSYRVTFAHHRILAYKOTLLPALMTLHTGWFPEENEGGVAASSQHTVLIINDNIAR ILGSAIVEDAKRKYHTALSTNSTATLHHAKAAYEORAGR"	
	complement(4349. 5134)	
gene	/gene="aur2f"	
	complement(4349. 5134)	
CDS	/gene="aur2f"	
	/note="Aur2f"	
	/codon_start=1	
	/transl_table=11	
	/product="ketoreductase-like protein"	
	/protein_id="AAK61717.1"	
	/db_xref="GI:14486276"	
	/translation="MSDQGRVAAVVGATSGICLAVARLLAVOGHRYFTGARSANVA ETYKELOGELEVDAVVDVRSSTGVAMIOAADRGSVDVNNAGRSQGVYADI ABEIMHDVIDTNLNSVFRREALTGTGLAKDGRGIINIASTAGKGVYAGPYAS KHCYGFETKALGNELAPGKITVNAVCCGYETPRAAQVROGXAAYATREAILLEKQ AKIPLGRISTPEEYAGLVGLIADSAASITISQALNVGGJGNF"	
	5386. 6657	
gene	/gene="aur2a"	
	5386. 6657	
CDS	/gene="aur2a"	
	/note="Aur2a"	
	/codon_start=1	
	/transl_table=11	
	/product="B-ketoacyl-ACP synthase-like protein"	
	/protein_id="AAK61718.1"	
	/db_xref="GI:14486277"	
	/translation="MSRRVYITGICVAVPGCVGTEKESLTLACKTATRAISLFDASP FRKRIAEADNPDLHGFTAEARLDRADQAFIVSAEAVRDSGIDTAELDPTRTGV TLGSAYCCTGLDVEYNAVSRGSGTMOVDHTLAVPHLEDVFPSSMAEVAVMVNAQG PVSWSYTGCTGLDSIGHAVELIREGSADVMTGTATAPISPIVACFDALKATSPRN DDPTASRPDPATNGFVLGCSAVLVELEDESARKGAHYATISGPAVSCNAYHMT GLRADGVEMAEATVVALDEARLDPDSADVYINAHSSGKQNDRHETAFKRSIGRAYE VPSSIKSMIGHSLGATGSSKVAASALAIETNIVPTANLHTPPDLCDLDDTTPLAKE QRTDAVLSVSGSGFGFGQSMILTRPRLGDA"	
	6654. 7916	
gene	/gene="aur2b"	
	6654. 7916	
CDS	/gene="aur2b"	
	/note="Aur2b"	
	/codon_start=1	
	/transl_table=11	
	/product="chain length factor-like protein"	

/protein\_id="AK61719.1"  
/db\_xref="GI:14486278"  
/translation="MTETAPARPGTPIPTTAPVYVIGVTPANGIGTPEAMAAVLR  
GSGCIRPVGREGATQYPAKAGEVPGPDAAENHPSRLPTDHTMTALTAADALAD  
AGVDPAELPDFSAGVITASSAGFEFGOKLQALMSGGQYVAISOFAPFVAVNIGQ  
ISTKNLGRPSGLVLEQAGLDVAQAKGKGLVYGVGVSSLCFPGWAHLA  
GGLSTSEDPARAFRPSAEADGHVAGGALVLEDAARAGAEIYGTAGYST  
FDPAPSGCAPRLQAAELALRDAGLTPADVVFADAAARDRAEARALTFGE  
YGVTPAPKTMGRLAAGASLDLAALSLRQVLPPTVNTGPAEDCDVDLVSEVR  
RPERLTAIVLWGRGCFNMAVYRAAH"

gene

CDS

7997..8500  
/gene="aur2C"  
7997..8500  
/gene="aur2C"  
/note="Aur2C: similar to acti dimerase from Streptomyces  
coelicolor"

/codon\_start=1  
/transl\_table=11  
/product="oxigenase-like protein"  
/protein\_id="AK61720.1"  
/db\_xref="GI:14486279"

gene

CDS

8524..9441  
/gene="aur2D"  
8524..9441  
/gene="aur2D"  
/note="Aur2D"  
/codon\_start=1  
/transl\_table=11  
/product="cyclase-like protein"  
/protein\_id="AK61721.1"  
/db\_xref="GI:14486280"

BASE COUNT

1370 a 3694 c 3132 g 1335 t

ORIGIN

Query Match

Best Local Similarity 54.1%; Pred. No. 1.3e-08;

Matches 216; Conservative 0; Mismatches 171; Indels 12; Gaps 2;

QY 365 taagatgcatgtcccaagacatacatctctctctgtggaactgagatgtgtc 424

DB 1526 TACGACGCGACCTCCCGCGCCGACCATCCGCGCGACCCAGGCGCGTAGAGTGTGTC 1467

QY 425 cgcctgtgaacggcgaagaacactctcccaacagcgtccacttgcaggtcttc 484

DB 1466 CG-----GCAGATCAACAGCTCCAGGTGAACCCCGGCTGCACCTGCACGGGCGCAC 1413

QY 485 tctcgaagctccctttagtgttggctgagacatacccaagcctggcgaataagat 544

DB 1412 GTGTCTCCGAGCGACGACGCGCTCCGATGACACCATCTCGCGCGCGGAGACGTACG 1353

QY 545 taactactaccccaagcagcgtgcgcgactgttgcgtacacatgacatgcatctcc 604

DB 1352 TACCGCTACCCGAAACACAGCCGCGCGCTGTGTGTCACGACACGACGCGCACAC 1293

QY 605 atcaaccgcaagcagcctcatatggtcagcgtgtgtctacatgacacagaccggct 664

DB 1292 CTGAGGCGCCAGACGCTTCTTGTGACGCGCTGTACCTCTCACCGACACACAC 1233

QY 665 gagatgacctgaacccctcccaagcagcgtacagcagttgatacccttggtctgact 724

DB 1232 GAGGCGACGCTGCGCTGCGGAGCGGTCC-----TACGACGTCCCGCTGTATCCGG 1179

QY 725 gccacgcatacaacgcagcagcactctcttccacc 763  
DB 1178 GACCGCGCGCTCGACGCGGACGCGACCCCTGCTCTACAC 1140

RESULT 15  
AP003233  
LOCUS

DEFINITION  
Oryza sativa genomic DNA, chromosome 1, PAC clone: P0037C04.

ACCESSION  
AP003233

VERSION  
AP003233.3

KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE



Mon Oct 7 11:28:36.2002

us-09-656-640a-1.rge

Page 18

[illegible]

Search completed: October 5, 2002, 20:23:09  
Job time: 8724 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 17:59:31 ; Search time 448.02 Seconds  
(without alignments)  
7503.501 Million cell updates/sec

Title: US-09-656-640a-1  
Perfect score: 1958  
Sequence: 1 ggcacccaacatgacgag.....ccgcctgagagtagaccggt 1958

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1958	100.0	1958	22	AAF30028
2	1956.4	99.9	1958	22	AAF82585
3	1946.2	99.4	2095	22	AAF82586
4	1946.2	99.4	2095	22	AAF82586
5	528.6	27.0	1959	14	AA047790
6	517.4	26.4	2110	14	AA261243
7	460.4	23.5	1791	20	AA227601
8	460.4	23.5	1791	21	AA550019
9	460.4	23.5	1791	21	AA551314

10	458.8	23.4	1791	20	AA225727	Stachybotrys chart
11	300.4	15.3	2067	20	AA225735	Stachybotrys chart
12	300.4	15.3	2067	20	AA227609	Stachybotrys pheno
13	300.4	15.3	3676	20	AA225728	Stachybotrys chart
14	300.4	15.3	3677	20	AA227602	Stachybotrys pheno
15	300.4	15.3	3677	21	AA550018	Stachybotrys chart
16	300.4	15.3	3677	21	AA551313	Stachybotrys chart
17	288	14.7	2063	21	AA550021	Curvularia pallasc
18	288	14.7	2063	21	AA551316	Curvularia pallasc
19	283	14.5	2905	21	AA550020	Bipolaris spicifera
20	283	14.5	2905	21	AA551315	Bipolaris spicifera
21	95.8	4.9	109519	22	AA508693	Microspora atrum
22	87.4	4.5	858	21	AA50022	Amerosporium atrum
23	87.4	4.5	858	21	AA51317	Amerosporium atrum
24	49.6	2.5	1533	22	AAH66018	C glutamicum codin
25	49.6	2.5	1614	22	AAH71220	Corynebacterium gl
26	49.6	2.5	349980	22	AAH68527	C glutamicum codin
27	47.2	2.4	523	21	AA43363	Rat secreted expre
28	47.2	2.4	523	21	AA43411	Rat secreted expre
29	44.8	2.3	10732	21	AA410594	Gene encoding a su
30	44	2.2	61313	23	AA559545	Propionibacterium
31	43.6	2.2	68356	22	AAK67283	Human Immune/haema
32	43.6	2.2	68356	22	AAK63212	Human Immune/haema
33	41.2	2.1	2012	22	AAH57385	Human skeletal mus
34	40.8	2.1	600	22	ABA59948	Human foetal liver
35	40.8	2.1	600	22	ABA28371	Probe #6837 for ge
36	40.8	2.1	600	22	AAK08219	Human brain expre
37	40.8	2.1	600	22	AAK34097	Human bone marrow
38	40.8	2.1	600	22	AAI39820	Probe #8506 used t
39	40.4	2.1	15772	22	AAK83220	Human Immune/haema
40	39.4	2.0	2071	15	AA058121	PH 2.5 acid phosph
41	39.4	2.0	2071	15	AA059645	A. niger pH 2.5 ac
42	38.8	2.0	523	21	AA481723	N. meningitidis pa
43	38.8	2.0	20757	20	AAK20599	Polynucleotide seq
44	38.8	2.0	23457	22	AAK74866	Human Immune/haema
45	38.8	2.0	31405	22	AAK67293	Human Immune/haema

#### ALIGNMENTS

RESULT 1	
ID AAF30028	AAF30028 standard; DNA; 1958 BP.
AC AAF30028;	
XX	
XX	
DT 23-APR-2001	(first entry)
XX	
DE	Stachybotrys chartarum phenol oxidase B spob gene.
XX	
KW	Phenol oxidising enzyme; phenol oxidase B; spob gene; bleach;
KW	pulp; paper; textile; detergent; ss.
XX	
OS	Stachybotrys chartarum.
XX	
FH	Key
FT	Location/Qualifiers
FT	13..1952
FT	/*tag= a
FT	1..203
FT	/note= "contains introns"
FT	/*tag= b
FT	204..255
FT	/*tag= c
FT	256..805
FT	/*tag= d
FT	806..888
FT	/*tag= e
FT	889..1628
FT	/*tag= f
FT	1629..1681
FT	/*tag= g
FT	1682..1958
FT	exon

FT /\*tag= h  
 PN US6168936-B1.  
 PD 02-JAN-2001.  
 XX  
 XX 22-SEP-1999; 99US-0401476.  
 XX  
 XX 22-SEP-1999; 99US-0401476.  
 XX  
 XX (GEMV ) GENENCOR INT INC.  
 PA  
 XX  
 XX Wang H;  
 PL  
 XX  
 XX WPI; 2001-136715/14.  
 DR P-PSDB; AAB20097.  
 DR  
 XX  
 XX New phenol oxidizing enzyme, also useful in the detergent, paper and  
 PT pulp, textile or food industries, especially in modifying the colour  
 PT associated with dyes and coloured compounds, as well as in anti-dye  
 PT transfer applications -  
 XX  
 XX Claim 5; Fig 1; 23pp; English.  
 XX

CC The present sequence is that of *Stachybotrys chartarum* M0C1.3898  
CC DNA encoding phenol oxidase B (see ABB20097). A full-length genomic  
CC sequence for the phenol oxidase B gene (spod) is given in AAF30029.  
CC The DNA was isolated by PCR using primers (see AAF30030-31) based  
CC on 2 peptide fragments of the isolated enzyme. The invention  
CC provides phenol oxidising enzymes such as phenol oxidase B, or  
CC enzymes with at least 68% identity to it, nucleic acids encoding  
CC them, expression vectors, filamentous fungus and yeast host cells,  
CC and methods for the recombinant production of the phenol oxidising  
CC enzymes. The enzymes are useful for bleaching of pulp and paper,  
CC fabric stains, and in detergent and textile applications. They  
CC show optimal activity at pH 5-11, 7-10.5 or 8-10, and 20-60 or  
CC 20-40 degree C.

Sequence 1958 BP; 394 A; 593 C; 490 G; 481 T; 0 other;

Query Match	100.0%;	Score 1958;	DB 22;	Length 1958;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1958; Conservative	0;	Mismatches	0;	Indels 0; Gaps

QY	1	ggatcatcaacatgatacgaacgaagcattatcgaaacgctgacctcgggacctggctgatac	60
Db	1	ggatccatcaacaagatacgaacgaagcattatcgaaacgctgacctcgggacctggctgatac	60
QY	61	ggcgcgcagctctctgtcgaatgcacagaatcccgctgtgtctgatacgaacgacatcgtcttcggt	120
Db	61	ggcgcgcagctctctgtcgaatgcacagaatcccgctgtgtctgatacgaacgacatcgtcttcggt	120
QY	121	ctcaaccagagagcagaagcagctgtagctccctccctcgtgacctgtacgaagtgtcctcgcgcg	180
Db	121	ctcaaccagagagcagaagcagctgtagctccctccctcgtgacctgtacgaagtgtcctcgcgcg	180
QY	181	atccctccctcgtgaagcgcgcgcacaaatgataagatactctatagctcgaagagccaaagt	240
Db	181	atccctccctcgtgaagcgcgcgcacaaatgataagatactctatagctcgaagagccaaagt	240
QY	241	tgcataatcatctgcagtacccgtccccaaccccaacaacatctggaagagacactctgtactaga	300
Db	241	tgcataatcatctgcagtacccgtccccaaccccaacaacatctggaagagacactctgtactaga	300
QY	301	gattgagagattagggcctctctcccaacgaatctacacctgatacttgagacgcgcacaatagt	360
Db	301	gattgagagattagggcctctctcccaacgaatctacacctgatacttgagacgcgcacaatagt	360
QY	361	tggatagagatgacatgccccacagacatacaatcatctgtcccgctgagcactaagagtg	420
Db	361	tggatagagatgacatgccccacagacatacaatcatctgtcccgctgagcactaagagtg	420

[illegible]

Db	961	tcaagctcttctgcgtctgtatcttgcctaccccttcgagagattcaagagacgaacttccctc	1020
Qy	1021	caggtcattctgcgccttgacagctgctgctgtctgtagaggccctgtgtgacaatgacactctgtac	1080
Db	1021	caggtcattctgcgccttgacagctgctgctgtctgtagaggccctgtgacaatgacactctgtac	1080
Qy	1081	atctctatggccgagcgctgtggagagtggttatcgaacttctcaactctgcctgtgccaatcc	1140
Db	1081	atctctatggccgagcgctgtggagagtggttatcgaacttctcaactctgcctgtgccaatcc	1140
Qy	1141	atcgatatccgacaacttccctgcgtgcctgcagacgctcctgcgtgtgtttagagccctgtgtataac	1200
Db	1141	atcgatatccgacaacttccctgcgtgcctgcagacgctcctgcgtgtgtttagagccctgtgtataac	1200
Qy	1201	actgacaaggtgatcgcgattcctgcctgtatgaaagctcctctgtagctgcgccgaactctctag	1260
Db	1201	actgacaaggtgatcgcgattcctgcctgtatgaaagctcctctgtagctgcgccgaactctctag	1260
Qy	1261	gtgcctgcacaactccgagatgctccttcccccagaggcggaacttggagaccccgcaaac	1320
Db	1261	gtgcctgcacaactccgagatgctccttcccccagaggcggaacttggagaccccgcaaac	1320
Qy	1321	cccactgttagccgagacttctcaacttccttcggcgtgcataaggagcagtggtgacaataacgga	1380
Db	1321	cccactgttagccgagacttctcaacttccttcggcgtgcataaggagcagtggtgacaataacgga	1380
Qy	1381	gttaacctctcgagatgtcgcgagaaacgctcctctccgaatgtgcccccgagacaactgttag	1440
Db	1381	gttaacctctcgagatgtcgcgagaaacgctcctctccgaatgtgcccccgagacaactgttag	1440
Qy	1441	atctgtgcgacttggaaacaactccaacgctgttgtagctacacccctgttacaatacactcgtt	1500
Db	1441	atctgtgcgacttggaaacaactccaacgctgttgtagctacacccctgttacaatacactcgtt	1500
Qy	1501	gaacttcgagatccttctcgttccaactgcgcgtgtgagatcgagccttataagagctgtcgtg	1560



OY	841	aagratcgatctatgaatattgtaoabgcttaacagtgctatcaagaagctgaagct	900
Db	841	aagratcgatctatgaatattgtaoabgcttaacagtgctatcaagaagctgaagct	900
OY	841	aagratcgatctatgaatattgtaoabgcttaacagtgctatcaagaagctgaagct	900
Db	841	aagratcgatctatgaatattgtaoabgcttaacagtgctatcaagaagctgaagct	900
OY	901	tggccatagctcaacgctgcagccgagaaftaacgccttcgcgtctcctcaacgctgcgc	960
Db	901	tggccatagctcaacgctgcagccgagaaftaacgccttcgcgtctcctcaacgctgcgc	960
OY	901	tggccatagctcaacgctgcagccgagaaftaacgccttcgcgtctcctcaacgctgcgc	960
Db	901	tggccatagctcaacgctgcagccgagaaftaacgccttcgcgtctcctcaacgctgcgc	960
OY	961	tcacgcctcttcgcgtctgctatcttgtaactctcgaagatctgaagaaacgaactctctc	1020
Db	961	tcacgcctcttcgcgtctgctatcttgtaactctcgaagatctgaagaaacgaactctctc	1020
OY	1021	caggttcacatgcgcgtctgaacggctggctcgtctgaagggccctggttgacaactgcgtac	1080
Db	1021	caggttcacatgcgcgtctgaacggctggctcgtctgaagggccctggttgacaactgcgtac	1080
OY	1081	atctctatagccgagcgctctggagaggttgttatcgactctccactctgcgtgcagatcc	1140
Db	1081	atctctatagccgagcgctctggagaggttgttatcgactctccactctgcgtgcagatcc	1140
OY	1141	atcgatatccgaaccccttcctgctgctgaagagctcgtgggttggttgagccctgagattgataac	1200
Db	1141	atcgatatccgaaccccttcctgctgctgaagagctcgtgggttggttgagccctgagattgataac	1200
OY	1201	acgtacaaggtctabgvcgaattcgttcgttgaabgaagtccttgagctgcgcgcgaacatctctgag	1260
Db	1201	acgtacaaggtctabgvcgaattcgttcgttgaabgaagtccttgagctgcgcgcgaacatctctgag	1260
OY	1261	gtgcgcctccgaacccctccgaagatgcttcctctcccggaagggcggaacattggaaaccccgcaaac	1320
Db	1261	gtgcgcctccgaacccctccgaagatgcttcctctcccggaagggcggaacattggaaaccccgcaaac	1320
OY	1321	cccaactatgcagagactcttcacactcttcgcgcgtgcgtcaatgagaaatgagaaatcaacaacga	1380
Db	1321	cccaactatgcagagactcttcacactcttcgcgcgtgcgtcaatgagaaatgagaaatcaacaacga	1380
OY	1381	gttaacctctcggatgtgcggagaaacgctctgctccgcaatgtgcgcgcgcgcgcgaacatgctgag	1440
Db	1381	gttaacctctcggatgtgcggagaaacgctctgctccgcaatgtgcgcgcgcgcgcgaacatgctgag	1440
OY	1441	atctcgcgcactctcgagatgctcggagaaacgctctgcgtccgcaatgagctcccgcgacactgtctgag	1500
Db	1441	atctcgcgcactctcgagatgctcggagaaacgctctgcgtccgcaatgagctcccgcgacactgtctgag	1500
OY	1501	gaattcccgagctccctctcgtctccactgcgcgtgcggctgagatcgagaccttatgagggctgctggtc	1560
Db	1501	gaattcccgagctccctctcgtctccactgcgcgtgcggctgagatcgagaccttatgagggctgctggtc	1560
OY	1561	ctcaagaagatgtgcgtgcgcgtgcgtctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt	1620
Db	1561	ctcaagaagatgtgcgtgcgcgtgcgtctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt	1620
OY	1621	cccttcgcgtaagctctcgcgtctcttaacctaactggtttctactcaatgcttaacatctaaac	1680
Db	1621	cccttcgcgtaagctctcgcgtctcttaacctaactggtttctactcaatgcttaacatctaaac	1680
OY	1681	gttggtgtctatcaatgttgacatgcgcaaacctgaatccacagagaaacagaagaaatgatagtgctg	1740
Db	1681	gttggtgtctatcaatgttgacatgcgcaaacctgaatccacagagaaacagaagaaatgatagtgctg	1740
OY	1741	ctttcaaatgttcaactggttctcgtgcgtacatgagctacacaactacacacgagttcatgtgacccca	1800
Db	1741	ctttcaaatgttcaactggttctcgtgcgtacatgagctacacaactacacacgagttcatgtgacccca	1800
OY	1801	tggagcctctctctggagagcccgccctctcctcctcgtgagaggttctgagaatgtgctcggtg	1860
Db	1801	tggagcctctctctggagagcccgccctctcctcctcgtgagaggttctgagaatgtgctcggtg	1860
OY	1861	acttcagcgagcgtctggccatcaactgacgcggaatccaagaagagtgctgacttcaaacccctacag	1920
Db	1861	acttcagcgagcgtctggccatcaactgacgcggaatccaagaagagtgctgacttcaaacccctacag	1920
OY	1921	cccaagcgtgaltgabtgcgcgcgtgagagatgaacccgt	1958

[illegible]

CC Applications.  
 XX Sequence 2095 BP; 437 A; 618 C; 510 G; 530 T; 0 other;

Query Match 99.4% Score 1946.2; DB 22; Length 2095;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1948; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 atccatcaaatgtatctgacgaagatctatccggtgtgctgtggtctgtgtatccg 62  
 DB 67 aacgcatcaaatgtatctgacgaagatctatccggtgtgctgtgtatccg 126  
 QY 63 cggagagctgtgcatgtacgaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 122  
 DB 127 cggagagctgtgcatgtacgaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 186  
 QY 123 caaccaagagcagaagcagctgaagtcctccctgtgtgtgtgtgtgtgtgtgtgt 182  
 DB 187 caaccaagagcagaagcagctgaagtcctccctgtgtgtgtgtgtgtgtgtgtgt 246  
 QY 183 cccctccctgaagcgcgcgaagtagtaagtatctatctatagtgtagcagaagcagcttg 242  
 DB 247 cccctccctgaagcgcgcgaagtagtaagtatctatctatagtgtagcagaagcagcttg 306  
 QY 243 ctatcatctgcatgacagctcccaaccacacactgtgagagagacatctgttactacagaga 302  
 DB 307 ctatcatctgcatgacagctcccaaccacacactgtgagagagacatctgttactacagaga 366  
 QY 303 tggagatgaagcctctctcccaacagatctacccctgtgtgtgtgtgtgtgtgtgtgt 362  
 DB 367 tggagatgaagcctctctcccaacagatctacccctgtgtgtgtgtgtgtgtgtgtgt 426  
 QY 363 gatcagatgtgcatgtctcccaagacatccatcatcgttccctcgtgtgtgtgtgtgtgtgt 422  
 DB 427 gatcagatgtgcatgtctcccaagacatccatcatcgttccctcgtgtgtgtgtgtgtgtgt 486  
 QY 423 tccgctctgtgacaagcgcgcgaagacacactctcccaacagcgttcacactgtgacagcctct 482  
 DB 487 tccgctctgtgacaagcgcgcgaagacacactctcccaacagcgttcacactgtgacagcctct 546  
 QY 483 tctctcagagctcctctgt 542  
 DB 547 tctctcagagctcctctgt 606  
 QY 543 attactactaccccaacagcgcgcgaagcgttcccgatgcttctgtgtgtgtgtgtgtgtgtgt 602  
 DB 607 attactactaccccaacagcgcgcgaagcgttcccgatgcttctgtgtgtgtgtgtgtgtgtgt 666  
 QY 603 ccatcaccgcgcgcgaagcgcgcgaagcgttcccgatgcttctgtgtgtgtgtgtgtgtgtgt 662  
 DB 667 ccatcaccgcgcgcgaagcgcgcgaagcgttcccgatgcttctgtgtgtgtgtgtgtgtgtgt 726  
 QY 663 ctgagagatgcccctgtacacctcccaacgcgcgcgaagcgttctgtgtgtgtgtgtgtgtgt 722  
 DB 727 ctgagagatgcccctgtacacctcccaacgcgcgcgaagcgttctgtgtgtgtgtgtgtgtgt 786  
 QY 723 ctgagagatgcccctgtacacctcccaacgcgcgcgaagcgttctgtgtgtgtgtgtgtgtgt 782  
 DB 787 ctgagagatgcccctgtacacctcccaacgcgcgcgaagcgttctgtgtgtgtgtgtgtgtgt 846  
 QY 783 tctgt 842  
 DB 847 tctgt 906  
 QY 843 gtatcagatgtatgaatgt 902  
 DB 907 gtatcagatgtatgaatgt 966  
 QY 903 gccatgtctcaagctgtgacgcgcgcgaagcgttcccgatgcttctgtgtgtgtgtgtgtgtgt 962  
 DB 967 gccatgtctcaagctgtgacgcgcgcgaagcgttcccgatgcttctgtgtgtgtgtgtgtgtgt 1026

QY 963 acgctcttcgctgtatctgtatctgtacactgtgagatctcaagaccagacttccctcca 1022  
 DB 1027 acgctcttcgctgtatctgtatctgtacactgtgagatctcaagaccagacttccctcca 1086  
 QY 1023 ggtcatgtgcgcgtgacgc 1082  
 DB 1087 ggtcatgtgcgcgtgacgc 1146  
 QY 1083 ctctatgtgc 1142  
 DB 1147 ctctatgtgc 1206  
 QY 1143 cgtatccgcgaacacttccctgt 1202  
 DB 1207 cgtatccgcgaacacttccctgt 1266  
 QY 1203 tgcacaagctcatcgatctgt 1262  
 DB 1267 tgcacaagctcatcgatctgt 1326  
 QY 1263 gccctgcgaacacttccctgt 1322  
 DB 1327 gccctgcgaacacttccctgt 1386  
 QY 1323 cactgtatgcagagacttccacttccgc 1382  
 DB 1387 cactgtatgcagagacttccacttccgc 1446  
 QY 1383 taacctctcgt 1442  
 DB 1447 taacctctcgt 1506  
 QY 1443 ctggc 1502  
 DB 1507 ctggc 1566  
 QY 1503 ctccgc 1562  
 DB 1567 ctccgc 1626  
 QY 1563 caagatgt 1622  
 DB 1627 caagatgt 1686  
 QY 1623 ttcccgtaagttctgc 1682  
 DB 1687 ttcccgtaagttctgc 1746  
 QY 1683 ggtgtctacatgt 1742  
 DB 1747 ggtgtctacatgt 1806  
 QY 1743 ttcaatgtacatgt 1802  
 DB 1807 ttcaatgtacatgt 1866  
 QY 1803 gaggcctctctgt 1862  
 DB 1867 gaggcctctctgt 1926  
 QY 1863 ttcaagcagactgtgcacatcagc 1922  
 DB 1927 ttcaagcagactgtgcacatcagc 1986  
 QY 1923 caagcgtatgtatgt 1953  
 DB 1987 caagcgtatgtatgt 2017

RESULT 4  
 AAF30029  
 ID AAF30029 standard; DNA; 2095 BP.  
 XX

AF30029; 23-APR-2001 (first entry)

Stachybotrys chartarum phenol oxidase B spob gene.

Phenol oxidizing enzyme: phenol oxidase B; spob gene; bleach; pulp; paper; textile; detergent; ss.

Stachybotrys chartarum.

Key	Location/Qualifiers
CDS	77..2016
FT	/*tag= a
FT	/note= "contains introns"
FT	1..267
FT	/*tag= b
FT	268..319
FT	/*tag= c
FT	320..869
FT	/*tag= d
FT	870..952
FT	/*tag= e
FT	953..1692
FT	/*tag= f
FT	1693..1745
FT	/*tag= g
FT	1746..2095
FT	/*tag= h

US6168936-B1.

02-JAN-2001.

22-SEP-1999; 99US-0401476.

22-SEP-1999; 99US-0401476.

(GEMV ) GENENCOR INT INC.

Wang H;

WPI; 2001-136715/14.

P-PSDB; AAB20097.

New phenol oxidizing enzyme, also useful in the detergent, paper and pulp, textile or food industries, especially in modifying the colour associated with dyed and coloured compounds, as well as in anti-dye transfer applications

Claim 5; Fig 3; 23pp; English.

The present sequence is that of the Stachybotrys chartarum MUC1 388998 spob gene encoding phenol oxidase B (see AAB20097). The gene was isolated by PCR using primers (see AAF30030-31) based on 2 peptide fragments of the isolated enzyme. The invention provides phenol oxidizing enzymes such as phenol oxidase B, or enzymes with at least 68% identity to it, nucleic acids encoding them, expression vectors, filamentous fungus and yeast host cells, and methods for the recombinant production of the phenol oxidising enzymes. The enzymes are useful for bleaching of pulp and paper, fabric stains, and in detergent and textile applications. They show optimal activity at pH 5-11, 7-10.5 or 8-10, and 20-60 or 20-40 degree C.

Sequence 2095 BP; 437 A; 618 C; 510 G; 530 T; 0 other;

	Query Match	99.4%	Score 1946.2	DB 22	Length 2095
	Best Local Similarity	99.8%	Pred. No. 0		
	Matches 1948: Conservative	0	Mismatches 3	Indels 0	Gaps 0
OY	3 atccatcaacatgatcgcgaagcatctcgagcgcgtgctctcgagccttgcgtatcgg				62

Db	67	aacgatacaactgtgtacagccaagctacatcggaagccggtgctctcgggcctgtgcgtgtatctcg	126
Qy	63	cggacagctctgtcgaatgcagatccggtctgtcgtgtcatctcgacagacatgctctccggtct	122
Db	127	cggcagcctctgcgaatgcagatccggtctgtcgtgtcatctcgacagacatgctctccggtct	186
Qy	123	caccaaaggcgacagacagcttgagttccctccctcgctctgtacgaagtgcctctgcgcgat	182
Db	187	caccaaaggcgacagacagcttgagttccctccctcgctctgtacgaagtgcctctgcgcgat	246
Qy	183	ccctctctctgaaggcgcccaagatgaatgatactctataggtctagccagagccaactgtg	242
Db	247	ccctccctctgaaggcgcccaagatgaatgatactctataggtctagccagagccaactgtg	306
Qy	243	ctaatacatgtgaatgaacggtctcccaacccaacactggaagagacatcttgtatactcagga	302
Db	307	ctaatacatgtgaatgaacggtctcccaacccaacactggaagagacatcttgtatactcagga	366
Qy	303	tggagatataggccctctctcccaacagatctacacccctatctctgagccgagccaactgtg	362
Db	367	tggagatataggccctctctcccaacagatctacacccctatctctgagccgagccaactgtg	426
Qy	363	gatacagatgacatgtctcccaaggacactacatcatgtctccctgtgtgcactgtgaagtgtg	422
Db	427	gatacagatgacatgtctcccaaggacactacatcatgtctccctgtgtgcactgtgaagtgtg	486
Qy	423	tcgcgtctctgaacagcgggagagaaacacccctctcccaagccgtctccactgtgacggtctct	482
Db	487	tcgcgtctctgaacagcgggagagaaacacccctctcccaagccgtctccactgtgacggtctct	546
Qy	483	tcctctcgagctccctctgtatgtgttggtctgtaggagacatacccaagccctgtgcgatacaag	542
Db	547	tcctctcgagctccctctgtatgtgttggtctgtaggagacatacccaagccctgtgcgatacaag	606
Qy	543	attactactacccaacaaaggcaggtcgcccgagatgctttgtgtacaaatgtagccaatgtagcgt	602
Db	607	attactactacccaacaaaggcaggtcgcccgagatgctttgtgtacaaatgtagccaatgtagcgt	666
Qy	603	ccatataccgcgcgaagaaacgacctatactgtgtctagaagctgtgtgtctatactgataccgaagcccg	662
Db	667	ccatataccgcgcgaagaaacgacctatactgtgtctagaagctgtgtgtgtctatactgataccgaagcccg	726
Qy	663	ctgagagatgacgttgaacacctccccaagcggtctacagcgagattgatatccctctgtgtctga	722
Db	727	ctgagagatgacgttgaacacctccccaagcggtctacagcgagattgatatccctctgtgtctga	786
Qy	723	ctgccaacacgataacaaacgacagacgcactctctctcccaacaaatgtagaggtgttccaaact	782
Db	787	ctgccaacacgataacaaacgacagacgcactctctctcccaacaaatgtagaggtgttccaaact	846
Qy	783	ctctgggtgtagcgttatctcaagttggttaagttagagcccatgtagatgctctcaagatccctagaa	842
Db	847	ctctgggtgtagcgttatctcaagttggttaagttagagcccatgtagatgctctcaagatccctagaa	906
Qy	843	gtatcagatgataatgtgacatgtgtcctctcaacagtgctataccaagaaaggtgcacgcttg	902
Db	907	gtatcagatgataatgtgacatgtgtcctctcaacagtgctataccaagaaaggtgcacgcttg	966
Qy	903	gactatgtctcaacgttgacgcgcgcgcaagataccgctctccgcttccctcaacgctgtccggtctc	962
Db	967	gactatgtctcaacgttgacgcgcgcgcaagataccgctctccgcttccctcaacgctgtccggtctc	1026
Qy	963	acgctctctgtccttgtatattgtctaacctcttgtagatctcaagagacagaattcccttcca	1022
Db	1027	acgctctctgtccttgtatattgtctaacctcttgtagatctcaagagacagaattcccttcca	1086
Qy	1023	ggtcaattccgctcaacgttgctcgtctctgtagggccctctgtgaacactgtgaacactgttaact	1082
Db	1087	ggtcaattccgctctcaacgttgctcgtctctgtagggccctctgtgaacactgtgaacactgttaact	1146
Qy	1083	ctctatagccagagcgtctgggaagttgttatacgactctccaaactctgcgtgtgcagatctcat	1142
Db	1147	ctctatagccagagcgtctgggaagttgttatacgactctccaaactctgcgtgtgcagatctcat	1206



[illegible]

Oy 1857 gaatgctcgggtgactcaagcagcttgcacactcgcgcattcagagatgctcag 1906  
 Db 1654 ggcctcagagtgccagctcaagtcgcttcagctgttactcagcgtattcagactatgctcga 1753  
 Oy 1907 ctctcaacctcagcccgactatga 1933  
 Db 1754 atacagaccttaacgcgcagactatcaga 1780

## RESULT

ID	AAZ61243 standard; DNA; 2110 BP.

AC AA

DT 30-MAY-2000 (first entry)

DNA encoding a phenol oxidising enzyme.

KW Phenol oxidizing enzyme; fungus; redox reaction; detergent; bleaching;

KW coloured compound; porphyrin; tannin; polyphenol; carotenoid;

XX

XX

FT	CDS	135..1943
----	-----	-----------

```

/product= "phenol oxidising enzyme"

```

PN WO200005349-A1.

PD 03-FEB-2000.  
VY

PF 13-JUL-1999; 99WO-EP04922.  
VY

PR 21-JUL-1998; 98EP-0202454.  
XX XX

PA (UNIT ) UNILEVER NV.  
PA (UNIT ) UNILEVER PIC

PA (HIND-) HINDUSTAN LEVER LTD.  
XY

PI	Convents D,	Gouka H
yy		

DR WPI; 2000-195101,  
DB P-PSDB: AAY69204

Phenol = oxidizing from Acromonium used in acetate cements

PT for bleaching stains on fabrics -  
xy

PS Example 2; Page 37-40; 45pp; English.  
 XY

CC The present sequence encodes a phenol oxidizing enzyme from the  
CC fungus *Ascomonium murorum*. The contig had the CBS accession number

CC 157.72: The enzyme catalyses redox reactions and is specific for molecular oxygen as the electron acceptor. The observed oxidising

CC is specifically used in detergents for bleaching strains on fabrics but also for bleaching pulp and for decolorization of

CC plant-derived food products. The enzyme has a pH optimum in the  
CC alkaline to neutral range and can hydrolyse a wide variety of coloured

CC compounds, e.g. porphyrins, tannins, polyphenols, carotenoids, anthocyanins and Maillard reaction products

Sequence 3110 BP: 400 A, 733 C, 556 C, 433 E, 0 other.

Query Match	26.48;	Score 517.4;	DB 21;	Length 2110;
Best Local Similarity	61.48;	Prod No 40-123;		

Matches 1041; Conservative 0; Mismatches 476; Indels 178; Gaps 6;

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84



Db 390 accgtcaaccaaccgttacaacggtgaggagatcgactactatgatagatcgatcaagac 449  
 QY 317 ttctcccaagatctactacgtgatcgagccggcacaactatggttgatcgatcgatg 376  
 Db 450 ttctctgagcaggtctctccctgacgtgaggcgtctgacccctcggtctcagacgacatg 509  
 QY 377 tccccaagactacatcatcgttctccgtggtcagctagatggttgcgtctcgttaac 436  
 Db 510 tcgcgcgtctactcttcatccagagagaaggcagggagagcgctgcgtctctgtaac 569  
 QY 437 agcggaagaagaacctctcccaacagcgtccacttgacaggtcttctctcgatccc 496  
 Db 570 a-----aggccactgtcagagctccatcatctcactcagcgtcggtccccc 623  
 QY 497 ttgatgttggtgagctgagacactacccagcgtcgagatgacagagttactactacccc 556  
 Db 624 tgggaagcgctggggtctgagatgtaacacccggcgagatcaaggacactactacacccc 683  
 QY 557 aacagcagagctgcccagctgtcttgatccatgacatgacatgacatccacgcgcag 616  
 Db 684 atgctgagcgagggcgctctctcgttgatccatgacatgacatgacatgacatgacgag 743  
 QY 617 aacgctacatggtcagctggtgtgtctacatgacagacccgctgagatgacctg 676  
 Db 744 atgcttacaacagcgagcgtgtgtgtgtacatcatccacgacctgagagatctctc 803  
 QY 677 aacctcccaagcgctacgagcaggttgatatacccttggtcttgatgacgacagatgac 736  
 Db 804 ggtctgcccagcgatatagtagatgtaacacatccctcactcactcagatcgaaacagatg 863  
 QY 737 aacgcagcagcactctctctctccacatgagaggtttccagctctggtggtgacgt 796  
 Db 864 aactccgacgtactctctctctcacaaggttgagagcagaggtctctggtggagatgct 923  
 QY 797 atccaagtgtagtgtagcccatgagatgctcagatccctcagaagatcatgatatga 856  
 Db 924 atccaagt----- 931  
 QY 857 aattgtgcatgcttaaccagtgctatcacagacggtcaagccttggtctatgctcaag 916  
 Db 932 -----gaagaggtcccggtgacctactctcagc 960  
 QY 917 tgcagccgcgaaatcacgcttcgcttccacagctgacgtctcagctctctcgtc 976  
 Db 961 ttgagcccaaggaatcacgcttcgctcgtctgagcgcgcgtctctcgtctcgtcgc- 1020  
 QY 977 tgtatcttgctactcctgagatcagaagacagactccctccaggtcatgtgcgtg 1036  
 Db 1021 tctactttgtcgaacccgcagatgagacacccgcacatcccggttcaggtatcgcttcg 1080  
 QY 1037 acggtgtcgtctgagggcccggttgacactgacactctgatactctatgctgagc 1096  
 Db 1081 acctgtgtctcctggagggagacccgctcagaccccaagctcgtctccctcactgtgagc 1140  
 QY 1097 gctggagaggtgttatcagcttccacactcgtggtgcagctcatcgatccagac 1156  
 Db 1141 gctacagatcatcctcagactctctcagactcagagcagacatcgagctccgcaag 1200  
 QY 1157 ttctgtgtgctgagcgtctggtgtgagcgtgagttgtatcactgacaagatgctgc 1216  
 Db 1201 agcccgcgctcgagcgtcgtgcatcgatgtaactacgacacagacgaagatgctgc 1260  
 QY 1217 gatt-----cgtcgttgatgaagctctgagtcgcccgaactctgaggtgctcgcaac 1273  
 Db 1261 gcttcaagtcgagcggagccctctctcgtccgcaacactcgtctgctccctccaccc 1320  
 QY 1274 tccgagatgttctctcccgagggcgcaactgaggaccccgcaaacccactgtagc 1333  
 Db 1321 tccgtgagctgacctctcc-----tccagcactctgacacacatcgacc 1365  
 QY 1334 agactttcaactcgcgcgtgataatggaagtggaacatcaacgagatctctcgtg 1393  
 Db 1366 actcgttcgcttcgcccagcgcgcgacgtgagatgacatcggtgtcatccttccg 1425

QY 1394 atgtcgaacacctgtctctccgaatgtgcccgcagacactggtgagatcgtgcgactg 1453  
 Db 1426 acgtcgaagaacgcctgtctgccaacggtgcccttgacacgcttcagctgtgagcagta 1485  
 QY 1454 agaaccaactccaacggttgagctacacactgttccatccactcgttgaactccgagtc 1513  
 Db 1486 cgaacgagcgagcgtcgtgagcatccatccatccatccatccatccatccatccatcc 1545  
 QY 1514 ttctc-----gttccactgcccgttgagctgagcgttatagagctg 1555  
 Db 1546 tgtctcgaacggtgtgtctacggtgtgtacgacccgtgtgtctgacgacgagctgtg 1605  
 QY 1556 ctggtccaagagatggtgtcgtgctgctgctgctgaggtgtgtatggttgagccact 1615  
 Db 1606 ctggtccgaagagcgtcgtctacactcgcgcaaggagagagcgtctcgtcgaagctact 1665  
 QY 1616 acgtccttcccgtaagttctcgtccttcaactggtttcactcatgatacatc 1675  
 Db 1666 acgtccttgccc----- 1679  
 QY 1676 tacaagtgtgtctacatgtgtgacgtgccacaacatgataccagagacacagatgat 1735  
 Db 1680 -----ggtgtgtacatgttccactgtccacaacatccatccacgagagacagatgat 1732  
 QY 1736 ggtcgttcaatgtacatgttctcgtgtgactatggttaacacacacagatgatgta 1795  
 Db 1733 ggtccgtctcaacgtctcagcgtccctccagactagcttacaacacacagcgtccctgta 1792  
 QY 1796 ccccatggaagcctctcgtgagggcccgccctctcctcctcgagagatgagaaatgctc 1855  
 Db 1793 ccccatgcaagagaggtctccgcgcagcagccgtagcttggaacaacagcgtcgagtgctcga 1852  
 QY 1856 ggtgtacttaacagagcttccatcactgacagcattcagagagatgctagcttcaacc 1915  
 Db 1853 gaagccttgaagcagcagagatcgagcgcgaaggtccagctgagtgacgactacgtcc 1912  
 QY 1916 ctacgcagagctga 1930  
 Db 1913 ctacgacaaccgga 1927

RESULT 7  
 AA27601  
 ID AA27601 standard; DNA; 1791 BP.  
 XX AC AA27601;  
 XX DT 16-DEC-1999 (first entry)  
 XX DE Stachybotrys phenol oxidase coding sequence.  
 XX KW Phenol oxidase; enzyme; coloured compound; dye transfer prevention;  
 XX KW fabric washing; stain bleaching; anti-dye transfer; detergent; ss.  
 XX OS Stachybotrys chartarum.  
 XX PN W09949020-A2.  
 XX PD 30-SEP-1999.  
 XX PF 23-MAR-1999; 99MO-US06327.  
 XX PR 24-MAR-1998; 98US-0046969.  
 XX PR 22-DEC-1998; 98US-0218702.  
 XX PR 22-MAR-1999; 99US-0273957.  
 XX PA (GENM ) GENENCOR INT INC.  
 XX PI Amory A, Wang H, Dhase P, Lambrechts-Kongvaux A, Wang C;  
 XX DR WPI; 1999-591088/50.  
 XX DR P-PSDB; AAY39992.

XX Novel enzyme for modifying coloured compounds used to prevent  
PT dye-transfer -  
XX  
PS Claim 21: Fig 5; 64pp: English.  
XX  
CC This sequence encodes the Stachybotrys chartarum phenol oxidase enzyme  
CC of the invention. The invention is used to modify a coloured compound and  
CC prevent dye transfer during fabric washing, or for stain bleaching or  
CC anti-dye transfer. It is useful in the detergent, paper and pulp, textile  
CC and food industries.  
XX  
SQ Sequence 1791 BP; 380 A; 551 C; 451 G; 409 T; 0 other;

Query Match 23.5% Score 460.4; DB 20; Length 1791;  
Best Local Similarity 59.3%; Pred.No.1.9e-116;  
Matches 996; Conservative 0; Mismatches 521; Indels 163; Gaps 6;

QY 257 accgtcccaacccaacacatgagagagacatctgtactacagatgagatagagcc 316  
DB 241 atattaccacaccccttcaacggcagagacattgtactatgagatcagatcaagcca 300  
QY 317 ttctccacacagatcaccttgatctgagacggccaacatggttgatagatgagatg 376  
DB 301 ttccagcaagaagattaccacaccccttgccctgacacatcgtcgtagatgagatg 360  
QY 377 tccccagagacacacacatcgtctcgtgagacagagatgtgttcgtcgtcgagac 436  
DB 361 agccctggcctcattcttcaatgttccagagagacagagatgagttaagttacatca 420  
QY 437 agcggagagac 496  
DB 421 a-----atgcacacgttgagagacatcgtacatcgtacacacacacacacacac 474  
QY 497 ttgtatgt 556  
DB 475 ttcgagtggt 534  
QY 557 aacagagacagac 616  
DB 535 aactac 594  
QY 617 aacgcctacatggttgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 676  
DB 595 aatgcctacttctgtcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 654  
QY 677 aacctccac 736  
DB 655 ggtcttcctcagtggt 714  
QY 737 aacgcagagac 796  
DB 715 aacgcgcagtgatcactcgt 774  
QY 797 attcaagtggttaagtgtgagccatgtgagatgtcttcatagtaagatcgtatga 856  
DB 775 atccatgtc----- 783  
QY 857 aattgtgcatgtcttcaacagtgctatacaagaagcgtacgctgtgctatgctcaag 916  
DB 784 -----aacgagacagcattgcttcttcaag 811  
QY 917 tgcagcgcgcagatcagccttcctcgtctcctcaacgctcgcgtccttctcgtc 976  
DB 812 tccagcccgccagatcagccttcccgatccctcaacgctcgcgtcgtcgtcgtcctcc 871  
QY 977 tgtatctgtacactcgtgagatcagagacacacacacacacacacacacacacacac 1036  
DB 872 tgaactcgtcagac 931  
QY 1037 acggt 1096

DB 932 atgctgtctccttcaagccccccttcaagacacacacacacacacacacacacacac 991  
QY 1097 gctggagagtggttatcagccttccacacacacacacacacacacacacacacacac 1156  
DB 992 gtacagagacatcattgacatcacaacacacacacacacacacacacacacacacac 1051  
QY 1157 ttctcgtcgtcagcgtctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1216  
DB 1052 ttgtcgaagac 1111  
QY 1217 gattcgt 1276  
DB 1112 gctcgtcgtcagccttgcacacacacacacacacacacacacacacacacacac 1168  
QY 1277 ggaatgttcccttcccccagagcgcgcaacacacacacacacacacacacacacacac 1336  
DB 1169 gtaacgttcccttccc-----tctcaagaagagcccgccgcaagaac 1213  
QY 1337 ctctcacccttcggcgtgtgtatagacagtgagacaaacacacacacacacacacacac 1396  
DB 1214 actcaagtttgaacgacagacagacacacacacacacacacacacacacacacacacac 1273  
QY 1397 tgcagaaacgctcgtcctcgcacatgtgtcccgacacacacacacacacacacacacac 1456  
DB 1274 tcaatagacgtgtcctcgtgcacacacacacacacacacacacacacacacacacacac 1333  
QY 1457 acaac 1516  
DB 1334 acccctcgtgagcgt 1393  
QY 1517 ctctgtcaccac 1573  
DB 1394 agcgaacgt 1453  
QY 1574 tctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1633  
DB 1454 tctgt 1508  
QY 1634 ttctcgtccttcaac 1693  
DB 1509 -----tgagatcaat 1520  
QY 1694 gttgcacatgcac 1753  
DB 1521 gttgcacatgcac 1580  
QY 1754 tgttctcgt 1813  
DB 1581 cgcacatgag 1640  
QY 1814 gaggcccgccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1873  
DB 1641 ggcgcgcgttcttcaac 1700  
QY 1874 tgcacatgcac 1933  
DB 1701 gtccatcactgc 1760

RESULT 8  
AAAS0019  
ID AAAS0019 standard; DNA; 1791 BP.  
XX  
XX  
AC AAAS0019;  
XX  
DT 10-OCT-2000 (first entry)  
XX  
DE Stachybotrys chartarum phenol oxidising enzyme cDNA.  
XX  
KW Phenol oxidising enzyme; detergent; bleaching; ss.  
XX  
OS Stachybotrys chartarum.  
XX

EH	Key	Location/Qualifiers
FT	CDS	7..1791
FT		/*tag= a
XX		
PN	MO200039306-A2.	
XX		
PD	06-JUL-2000.	
XX		
PF	20-DEC-1999;	99WO-EP10287.
XX		
PR	23-DEC-1998;	98US-0220871.
XX		
PR	23-JUN-1999;	99US-0338723.
XX		
PA	(UNITL ) UNILEVER NV.	
XX		
PA	(UNITL ) UNILEVER PLC.	
XX		
PI	(HIND-) HINDUSTAN LEYER LTD.	
XX		
PI	Bodie EA, Van Der Velden S, De Vries CH, Wang H;	
XX		
DR	WPI: 2000-514528/46.	
XX		
DR	P-PSDB: AAY95537.	
XX		
PT	Detergent composition comprising novel phenol oxidising enzyme obtained	
PT	from fungus or bacteria, useful for pulp and paper bleaching, bleaching	
XX	color of stains on fabric and for anti-dye redeposition	
XX		
PS	Disclosure: Fig 5A-B: 45pp; English.	
XX		
CC	The present sequence is that of the Stachybotrys chartarum MUC1 38898	
CC	cDNA encoding a phenol oxidising enzyme (see AAY95537). The invention	
CC	relates to detergent compositions comprising novel phenol oxidising	
CC	enzymes that are encoded by nucleic acids capable of hybridising to	
CC	the S. chartarum phenol oxidising enzyme gene (see AA50018), provided	
CC	the enzymes are capable of modifying the colour associated with dyes	
CC	or coloured compounds, and are produced from a bacterium, yeast or	
CC	fungus (see AAY95538-40). The phenol oxidising enzymes can be used	
CC	for pulp and paper bleaching, for bleaching the colour of stains on	
CC	fabric and for anti-dye transfer in detergent and textile	
CC	applications. They may also be capable of modifying the colour in	
CC	the absence or presence of an enhancer. Expression vectors and host	
CC	cells comprising a nucleic acid encoding a phenol oxidising enzyme,	
CC	methods for producing the phenol oxidising enzyme, and methods for	
CC	constructing expression hosts are provided.	
XX		
XX		
SQ	Sequence 1791 BP; 380 A; 551 C; 451 G; 409 T; 0 other;	
	Query Match	23.5%; Score 460.4; DB 21; Length 1791;
	Best Local Similarity	59.3%; Pred. No. 1.9e-116;
	Matches 996; Conservative	0; Mismatches 521; Indels 163; Gaps 6;

DB	535	aactaacatccgcgcgccttcgtgtgtatccatgacacacgcttcctcatgaagactgtgag	594
QY	617	aagcctacatgtgtgaagctgtgtgtatcatatccagacacccgctgagatgcctg	676
DB	595	aatgcctactctgtcagcgtcgtgcgcctacatcatcaacagcagagctgagatgcctc	654
QY	677	aacctccacgagcgtacagcgtgtgtatccacctgtgttcagatcccaagcagctac	736
DB	655	ggtcctcctatgtgtgctctgtgcagatcgcatacctcctgtcctgtacagcgaactat	714
QY	737	aacgcagacgac	796
DB	715	aacgcagatgtatccctgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgc	774
QY	797	atccagatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtat	856
DB	775	atccatgtc-----	783
QY	857	aattgtcatgtctaacacagctgtatcacagacggttcagcctgtgctatgtcagc	916
DB	784	-----aacgcagacgcatgtgcttccttaag	811
QY	917	tgaagccgcgaagtaacagccttcctgcctcctcctcctcctcctcctcctcctcctc	976
DB	812	tccagccgcgaagtaacagccttcctgcctcctcctcctcctcctcctcctcctcctc	871
QY	977	tgtatctgtac	1036
DB	872	tctacacgtcgaagac	931
QY	1037	acggtgtcgtcgt	1096
DB	932	atgt	991
QY	1097	gtcgtgaggt	1156
DB	992	gtacagagatcatatgtatgtatcaacacacacacacacacacacacacacacacacacac	1051
QY	1157	tctcgt	1216
DB	1052	tgtcgtgagac	1111
QY	1217	gattcgtcgt	1276
DB	1112	gtctgt	1168
QY	1277	gagatgtctccttcccgagggcggaactgtggaccccgcaaaccccaactgtatgcaga	1336
DB	1169	gtgagcttcttcc-----tcccaagaagggcccgcaagaac	1213
QY	1337	cttccacttcggt	1396
DB	1214	acttcaagtttgaac	1273
QY	1397	tctgagacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	1456
DB	1274	tcaatagagctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	1333
QY	1457	acaactcaacaggttgaactac	1516
DB	1334	accctccttgtaggt	1393
QY	1517	ctcgttccactgcgcgt	1573
DB	1394	agcgaactgt	1453
QY	1574	tctggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1633
DB	1454	tctggttggcaggggt	1508
QY	1634	tctcgtccttctac	1693
DB	1509	-----tgagacttaacat	1520





```
Db 872 tctacctctgcaagaccagctctcccaagctcacaagcttcccttcccaagctcatgtcctctg 931
QY 1037 acggttgctcgtcttgaggccctgttgacactgacactgtacatctctatgtccgagc 1096
Db 932 atgctcgtccctcaagcccccgtctcagaactctaaccttcaactgtgctgtgagc 991
QY 1097 gctggaggtgtgtatcatgactctccacctctcgtcgcagctccatcgatccagacc 1156
Db 992 gttacagagatcatattatgacttcaacaaacttctgtgcagactcttgactctgagcag 1051
QY 1157 ttccctgtctgacagtgctcgtgtgtgagccttgagcttgatgaactgacaagtgatgc 1216
Db 1052 ttgctgagacaacagatgctgcagcagagatgagtagctgcagactctcagagtgtatgc 1111
QY 1217 gattcgtcgttgatgaagtctcttgagtcgcgcagcactcttgaggtgtgcctgcaactcc 1276
Db 1112 gcttcgtcgtcagcagctgtgacgtgtgag---gacacaagcagagtcctccctccctcc 1168
QY 1277 gagaatgtctcttcccccagagcgcgcaactgagagcccgcaaaacccactgtatgacaga 1336
Db 1169 gtgagcttctcttcc-----tctcacaaggaagcgcccgccgacagaagc 1213
QY 1337 cttcaccttcgcgcgtgtgacatgagacagtggacaatcaacggagatcaactctcgatg 1386
Db 1214 actcaagatttgacagcagacagacgacactaacctgacatgactgtgtgttgcgagtg 1273
QY 1397 tggagaaacgtctgctcgcgaatgtgcccgcgacactgttgaagtgtgagcttgagcttga 1456
Db 1274 tcaatgagcgtctcgtcgcagacagccgagctgtgacacgttgaagcttgagagctcgaga 1333
QY 1457 acaactccaacgggttgagactaacctgttacaattcaactcgttgacttcgagctctt 1516
Db 1334 accctctgagagctgagcagcccgctccacattcaactgtgtgacttcaagctctca 1393
QY 1517 ctgtctcactgacccgag---agtcagagcctatgagctgtcgtcgtcgaagatgttg 1573
Db 1394 agcgaactgtgtgctcgtgcagagctcagctcagctacagctgtcgtcttaagatgtcg 1453
QY 1574 tctgctcgtcgtcgtcgtgagaggtgtgctatgttgagagcccaactacgtccttccgtaag 1633
Db 1454 tctggttgagcagaggtgagagaccttgacatcgagagcccaactcaacctgtgagc----- 1508
QY 1634 ttctgccttcttaactaactcgtgttctcaatcagctacatcaagtggtgtctacat 1693
Db 1509 -----tgtagcttcat 1520
QY 1694 gttgcaactgcaacaactgtatccagagagacgaacatgagctgttcaatgacac 1753
Db 1521 gtggcaactgcaacaactcattcaagagabaagacatgagctgtatcaacgctcac 1580
QY 1754 tgttctcgtgactatgctacaactacacagatctcatgtgaccccaatggagcctctctg 1813
Db 1581 cgcacatgagagagaagatatcttcaagagagactcgtgagagcccaatgaaacccaagt 1640
QY 1814 gaagcccgccctctcctcgtcgtgagaggtcgagaatgctcgggtgacttcaagagact 1873
Db 1641 gcgcgcgcttctcaacaacccaacgacttccatgctcgcgtggaacttctcgcgcga 1700
QY 1874 tggcatcactgacacgcatatgagagatgctgtagcttcaaacccctacagagtgtaga 1933
Db 1701 gtccatcaactgcgcagatgacagagctgcgcgagcagagacgctlaaacgcgctcgatga 1760
```

RESULT 11

AA25735 standard; DNA: 2067 BP.

```
XX AA25735;
XX 05-JAN-2000 (first entry)
DT Stachybotrys chartarum phenol oxidising enzyme PCR fragment.
DE
XX
```

```
KM Stachybotrys chartarum; phenol oxidising enzyme; colour: dye;
KW detergent; anti-dye transfer; stain removal; bleaching; PCR primer; ss.
OS Synthetic.
OS Stachybotrys chartarum.
XX MO9949010-A2.
XX PD 30-SEP-1999.
XX PF 23-MAR-1999; 99MO-EP02042.
XX PR 24-MAR-1998; 98US-0046969.
XX PR 22-DEC-1998; 98US-0218702.
XX PA (UNITL ) UNILEVER NV.
XX PA (UNITL ) UNILEVER PLC.
XX PI Convents D, Amory A, Wang H, Dhasee P, Lambrechts-Rongvaux A;
XX PI Wang C;
XX DR WPI; 1999-601211/51.
XX PT Detergent composition containing phenol oxidase from Stachybotrys, used
XX PT to bleach stains and prevent dye transfer
XX PS Example 17; Fig 9; 56pp; English.
XX CC The present invention describes a detergent composition containing a
XX CC purified phenol oxidising enzyme derived from Stachybotrys. The present
XX CC sequence represents a PCR fragment of Stachybotrys chartarum phenol
XX CC oxidising enzyme. The enzyme can be used to modify the colour of dyes
XX CC and other coloured compounds (e.g. for use in pulp and paper bleaching
XX CC also for removing stains; e.g. food, tea, blood etc., from fabrics) and
XX CC for preventing dye transfer during fabric washing.
SQ Sequence 2067 BP; 445 A; 621 C; 500 G; 501 T; 0 other;

Query Match 15.3%; Score 300.4; DB 20; Length 2067;
Best Local Similarity 55.28; Pred. No. 2.4e-72;
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

QY 316 ctctcccaacatctacactgtatctgagccgccaacatgttgatagatggat 375
Db 457 ctgaccccttcagattacaccacactgtgcctgcacactctgtgcgtaagatggat 516
QY 376 gtcccaagacatacatgttctcgtgacatgagatgtgtccgtctgtga 435
Db 517 gaagccctgtctacttcaatgttccagagagaacagagactgttagtgcacaa 576
QY 436 cagcagagagaacacactctcccaacagctcacttccagagccttctcgaagctc 495
Db 577 ca-----atgcacagtgagaactcgttccatctgcagagctcccatcggtgcc 630
QY 496 ctgtatgtgttggtctgagacataccacacagctgagagtaagaatatactacc 555
Db 631 ttgcagtgtgtggtcgaagatgtgacacttccctgtgagtagtaagaattactacttcc 690
QY 556 caacagcagagctgcgcgcatgtcttgatccatgacatgc----- 597
Db 691 caactacaatcgcgcgctctctgtgtacatgacacagccttcaatgaagtagtcta 750
QY 598 -----catgtccatcacgcgcgag 616
Db 751 cgagccttactcttctgtactacttgtgtaacaactcttctgtatgactgtgag 810
QY 617 aacgcctacatggtcagagctgtgtgtacatgatacgaagccgctgagatgacctg 676
Db 811 aatgcactacttgttcagagctgtgcgtacatcatcaacagagagctgtgagatgcttc 870
QY 677 aacctccagcgctacagcgagttgatatacccttgttctgactgccaagcgatcac 736
```

```

Db      871 ggtcttctactagtgatgagatgcatatccctcttgatctctgagccaaagtactat 930
Qy      737 aacgcagacgcacactctctctccaccacatgagagttccacgtcttggtgacgtt 796
      931 aacgcagatggtaccctcgctctgacccgagggagagacacagacactggtggagagatgc 990
Qy      797 attcaagtgttaagtgtgagcccatgtgagatgtcagatccctagaagaatcatgatatga 856
      991 atccatgtc----- 999
Qy      857 aattgtagatgctcttaaccagtgctatcacagaaagtcacgtcttgacctatgctcaag 916
      1000 -----aacgagacagacatgagccttctcttcaag 1027
Qy      917 tgcagccgcagaaftaacggtctccgctctcctcaacgctgacgtctcactctctgctc 976
      1028 tccagcccccagaaftaacggtctcctcctcaacgctgacgtctcctgctctgctctc 1087
Db      977 tcatcttgctactaccttgagagatcagagacacagactctcctcccaagttacattgcgctg 1036
      1088 tctaccctgcagagacagactctcccaagctcagaattctcttcccaagttacattgcctctg 1147
Qy      1037 acggtgtctgctgtgagggccctggtgacactgacactctgtacatctctcatgcccagac 1096
      1148 atgctgtctcctcctcaagcccccgttcagacacttaaccttaacctctgtctgtccgagc 1207
Qy      1097 gctggagagttgtta----- 1111
      1208 gttacagagatcattatgtatgctcctccctcctcaacgaatgagtcagaactcctaagac 1267
Qy      1112 -----tcgactctcctcctctgctgagtcagttacatgatalccgaacctctc 1162
      1268 taacactgtgagactctcaccacacttgctgcagactcttgacactgcgacagctgtctg 1327
Db      1163 gtctcagcgtctcgtgtgtgagcgtgagtttataaacactgacagagtcagtcagatcg 1222
      1328 agaccacagatgctcgagacagatgtagctcgcacactctcgcagtgatgctgcctcg 1387
Qy      1223 tctgttagagtcctgtgagtcgcccacactctcgtgagtcgagtcgacacactccagatg 1282
      1388 tctcagactctgacactggtgag-----gacaaacagacagtcctcctcactcccgtagc 1444
Qy      1283 ttccttccccgagggcgcaactgtagacccgcgaacccccactgtagacagacttca 1342
      1445 ttccttccct-----cctcaagaagggcccgccgacagaacttca 1489
Db      1343 ccttgcgcgtgtctatagtcacagtgtgacaatcaacgagttacactctcgatgtcgaga 1402
      1490 agtttgaacgcagacagagacactacactgtaacgagtgtgtgcttgcgcatgtcatg 1549
Qy      1403 accgtctgcgcgaatgtgccccgcagacactgttgagatctgagactgagaaacaact 1462
      1550 agcgttctcgcgcaagcccgagactcggacccgttggagtcctggagactcgagaacccct 1609
Db      1463 ccaacggttggactcaaccctgttacaatcaactcgttgaactcgcagtccttctcgt 1522
      1610 ctgagagctgtagacacaccgcctccacatcacttgcttgcttcaagaatcctaagagaa 1669
Qy      1523 ccactgcgcgtg-----agtgagacctatagagctgctgtgtctcaagaagtgtctg 1576
      1670 ctgtgtgtctgagcagagttacccctacagatctgtgtcttaagagatgctcgtcgt 1729
Db      1580 tggctcgtctgagagttgtctatgttagagccactacgctccttcccgtaagttctg 1639
      1730 tgggacggggtgagacccctgacatcgagggccacatcaaccctgagc----- 1778
Qy      1640 cctttaactaactggttttcaactcatgtacaatctacaagtgtgtctacatgttga 1699
      1779 -----tgagacttaaatgttga 1796
Db      1700 ctgcacacactgatacagagacacagacatgagtcgcttcaatgttaactgtctt 1759
      1797 ctgtcaaacatcttcaacgagataacgacatgagctgtatctcaacgctcacacgcat 1856

```

```

Qy      1760 cgttgactatggtacactacacacgagttcatgtaccccatgagcctctcgagcc 1819
      1857 gggaggaagaagatatcttcaagagagacttcgagaccatgaaacccaagtgggcgcc 1916
Db      1820 ccgcctctctctcctcgagagatcgagaatggtcgcgtgtaacttaacgagacttgcacat 1879
      1917 cgttcttaacacgcagacactcatcatgctgcgttggaacttctcgcgcgagctcat 1976
Qy      1880 cactgcacgcatcagagatgtagcttcaacccctacagcccaagctgatatga 1933
      1977 cactgcgcagtgatagagagctgagccgagagagcgttaacacgcctcatga 2030
Db

RESULT 12
AAZ27609
ID  AAZ27609 standard; DNA: 2067 BP.
XX
XX  AAZ27609;
AC  AAZ27609;
XX  16-DEC-1999 (first entry)
DT
DT  Stachybotrys phenol oxidase coding sequence.
DE
XX  Phenol oxidase; enzyme; coloured compound; dye transfer prevention;
KW  fabric washing; stain bleaching; anti-dye transfer; detergent; ss.
XX
XX  Stachybotrys chartarum.
OS
XX  WO949020-A2.
PN  30-SEP-1999.
XX
XX  23-MAR-1999; 99NC-US06327.
PF
XX  24-MAR-1998; 980S-0046969.
PR  22-DEC-1998; 98US-0218702.
PR  22-MAR-1999; 99US-0273957.
XX
XX  (GENM ) GENENCOR INT INC.
PA
PI  Amory A, Wang H, Dhase P, Lambrechts-Rongvaux A, Wang C;
PI  WPI; 1999-591088/50.
DR
XX  Novel enzyme for modifying coloured compounds used to prevent
PT  dye-transfer -
PS
PS  Example 17; Fig 9; 64pp; English.
XX
XX  This sequence encodes the Stachybotrys chartarum phenol oxidase enzyme
CC  of the invention. The invention is used to modify a coloured compound and
CC  prevent dye transfer during fabric washing, or for stain bleaching or
CC  anti-dye transfer. It is useful in the detergent, paper and pulp, textile
CC  and food industries.
XX
XX  Sequence 2067 BP; 445 A; 621 C; 500 G; 501 T; 0 other;
SO

Query Match      15.3%; Score 300.4; DB 20; Length 2067;
Best Local Similarity 55.2%; Pred. No. 2,4e-72;
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;
Qy      316 ctctccacacagatcaccttgatctcgagagccgcgaacatggttggataagatgcat 375
      457 ctgaccccttcagatthtaacccacacttgccctgcactctcctgctggtacagatgcat 516
Db      376 gtccccagacactacatcatgctctcgttgagacatgagatgttgcgcctgtgaa 435
      517 gagccctggtccacttaattgaatgttccagagagaaacagacagatgtatgtatcaaa 576
Qy      436 cagcggagagaacacactctcccaacagcgtccacacttgacagctcttctcagagctcc 495
      495

```

Db 577 ca-----atgccacgcgtggagaactcgtccatctgacagcgtccccaatcgctgccc 630  
 QY 496 ctttgatgcttgagcgtgaaggacaactaccagcgtcgagtaacaaggattactactccc 555  
 Db 631 ttccgaatgcttgagcgtgaaggacaactaccagcgtcgagtaacaaggattactactccc 690  
 QY 556 caacagcagcgtcgagcgtgaaggacaactaccagcgtcgagtaacaaggattactactccc 597  
 Db 691 caactaccatccgcgcgtcgtgtaacatgaccagcgtccttaagatgacta 750  
 QY 598 -----catgtccatcaacgcgcag 616  
 Db 751 cgaagccttaactctcttgctacacttggttaacacactcttctgtagatgctgag 810  
 QY 617 aagcctacatgagtgacgctggtgtatcatgataccagacccggttgagatgctg 676  
 Db 811 aagcctacatgagtgacgctggtgtatcatgataccagacccggttgagatgctg 870  
 QY 677 aacgtccacagcgtgacggtgagttgataccacttggtctgactgccaagcgaatac 736  
 Db 871 ggtcttcctagtgactaaggcaggttcgataccctctgactgccaagcgaatac 930  
 QY 737 aagcgaacagcgtcgtctctctccacaaatgagaggtttccacgctcttggtgagct 796  
 Db 931 aacgcgaatgacccgtcgctccgaaggtgagagacgaagacccgttggtgagatgct 990  
 QY 797 attcaagtggttaagttgagccatitgagatgctcagatccataagtatcgatgata 856  
 Db 991 atccatgtc----- 999  
 QY 857 aattgtgacatgcttaacacagtgctatacaagaaggtcaagccttggtcctatgctcaag 916  
 Db 1000 -----aagcagacagcagcagcgtcttcccttaag 1027  
 QY 917 tgcagccgcgaagtacagcttccgcttccctcaacagcgtcgatcgaagcgtctctgctc 976  
 Db 1028 tccagccgcgaagtacagcttccgcttccctcaacagcgtcgatcgaagcgtctctgctc 1087  
 QY 977 tgaatcttgctacactctgagagatcagaagacagacttccctccaggtatgacgtg 1036  
 Db 1088 tgaatcttgctacactctgagagatcagaagacagacttccctccaggtatgacgtg 1147  
 QY 1037 aaggttgctgcttgagagcgttggtgacacgtgacactgctgatactctatgacgagc 1096  
 Db 1148 atgctgcttccctcaagcccgcttcaagacttaacacttactgctggttgacgagc 1207  
 QY 1097 gctcgagaggtgtta----- 1111  
 Db 1208 gttacagagatcaltatgtgatatgccccctccctccacgaatgagtaagaacttaagac 1267  
 QY 1112 -----tcgaacttccacacttgcgttgccagttccatcgatataccgaacttctg 1162  
 Db 1268 taacacttgtagacttccacacttgcgttgccagacttctgacttgccagacttgctg 1327  
 QY 1163 gttcgtgacggttcgtggtgagcgtggtgataaactgacagaggtatgactgactg 1222  
 Db 1328 agaccacaagatgctcgagagatgtagactgacgtcgtcgtgaggtgactgactg 1387  
 QY 1223 tctgttagatgaggtcgtgagtcgccccacacttctgaggtggtggtgacacactcgagatg 1282  
 Db 1388 tctgttagatgaggtcgtgagtcgccccacacttctgaggtggtggtgacacactcgagatg 1444  
 QY 1283 ttccttcccccagggcggaacttggaacccgcaaaccccaactgtagagagacttca 1342  
 Db 1445 ttccttccccc-----cctcaagaagagcccgccgacacaagacttca 1489  
 QY 1343 ccttgagcgttgatagacagatggaacaaacagagatgaactctcgatgctgaga 1402  
 Db 1490 agtttagacgcagacaagagacactgatacagagatggttggttgcgactgactg 1549  
 QY 1403 accgtctgctgcgaatggtgccccgcgacactgttgagatctgagcgttgagaacact 1462  
 Db 1550 agcgttctgctgcgaatggtgccccgcgacactgttgagatctgagcgttgagaacact 1609

QY 1463 ccaacggttgagtaacactcgttcaactacactcgttgactcgttccttctcgtt 1522  
 Db 1610 ctgagagcgttgagtaacactcgttcaactacactcgttgactcgttccttctcgtt 1669  
 QY 1523 ccaatgcccgttgg---agtcgaactatgaggtcgttgcttcaagagatgtgtcgtg 1579  
 Db 1670 ctggtgtgcttgagcaggtatgcccatacgaactcgttgcttcaagagatgtgtcgtg 1729  
 QY 1580 tggctgctgaggtgtgtatggttgaggtccacacacgtccttccgttaagttctg 1639  
 Db 1730 tggcgaaggttgagacccctgacatccaggtccacacacgtccttccgttaagttctg 1778  
 QY 1640 ccttaccatacgtgttccatcagatcgaatcgaatgaggtgtgtacatgttga 1699  
 Db 1779 -----tggagcttaactgtggca 1796  
 QY 1700 ctgcacaacactatccacagagacacagacatgagctgtcttcaatgtcactgttct 1759  
 Db 1797 ctgacacaacactatccacagagacacagacatgagctgtcttcaatgtcactgttct 1856  
 QY 1760 cgtgtgactatggttacaacacacaggttcattgaccccatgagcgtctctggagc 1819  
 Db 1857 ggaaggaaggtatatacttcaagagagacgttcgaagcccatgaaacccaatgagcgc 1916  
 QY 1820 ccgccccttctcctcgtcgagagatcgagaaatggtcgtcggtgacttcaagcagcttgc 1879  
 Db 1917 cgttcttcaacacagcgaacgttccatgctgctgctgagaaacttctccgcgagttcat 1976  
 QY 1880 cactgacgcgattcagagagatggtgacttcaacccctcaagcccaagctgata 1933  
 Db 1977 cactgacgcgattcagagagatggtgacttcaacccctcaagcccaagctgata 2030

## RESULT 13

AAZ25728  
ID AAZ25728 standard; DNA: 3676 BP.

XX AC AAZ25728;

XX DT 05-JAN-2000 (first entry)

XX XX Stachybotrys chartarum phenol oxidising enzyme genomic DNA.

XX XX Stachybotrys chartarum; phenol oxidising enzyme; colour; dye;

KW detergent; anti-dye transfer; stain removal; bleaching; ds.

XX OS Stachybotrys chartarum.

XX PN WO949010-A2.

XX PD 30-SEP-1999.

XX PF 23-MAR-1999; 99WC-EP02042.

XX PR 24-MAR-1998; 98US-0046969.

XX PR 22-DEC-1998; 98US-0218702.

XX PA (UNITL ) UNILEVER NV.

XX PI (UNITL ) UNILEVER PLC.

XX PI Convents D, Amory A, Wang H, Dhasee P, Lambrechts-Rongvaux A;

XX DR WPI: 1999-601211/51.

XX DR P-PSDB; AAY45222.

XX PT Detergent composition containing phenol oxidase from Stachybotrys, used

XX PS to bleach stains and prevent dye transfer

XX CC Example 14; Fig 6; 56pp; English.

The present invention describes a detergent composition containing a



CC purified phenol oxidising enzyme derived from *Stachybotrys*. The present  
CC sequence encodes *Stachybotrys chartarum* phenol oxidising enzyme. The  
CC enzyme can be used to modify the colour of dyes and other coloured  
CC compounds (e.g. for use in pulp and paper bleaching also for removing  
CC stains, e.g. food, tea, blood etc., from fabrics) and for preventing dye  
CC transfer during fabric washing.

XX Sequence 3676 BP; 822 A; 1056 C; 849 G; 948 T; 1 other;

Query Match 15.3%; Score 300.4; DB 20; Length 3676;  
Best Local Similarity 55.2%; Pred. No. 3.2e-72;  
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

```

QY 316 ctctcccaacatattacccatgattgtgagcgagcaaatgtgttgatagatggat 375
DB 1491 ctgacccttcaatltaccaccccttgcgcctgcacactctcgtcggtacagatggat 1550
QY 376 gtcccaagagacatcatatcgtctcgtggaactgagtggtgtccgctgtgaa 435
DB 1551 gaggcctgtctacttcaatgattccagagagaaacagactgttagtgatcaca 1610
QY 436 cagcgagagaaacactctcccaacagcgttcaacttgacagctcttctctcgaagtc 495
DB 1611 ca-----atgcacacgtgagaaactcgtccactcgtcacggtcccaatgcgtgcc 1664
QY 496 ctgtgattgttggtgagagacactcccaagcctgagagataagattactacc 555
DB 1665 ttccgattgttggtgagagaaatgtgaccttccctgagagataagattactacc 1724
QY 556 caacagagagcggtccgcgtctgtgtgatacctgacacacac----- 597
DB 1725 caactaacatccgcgcgcctctgtgtgatacctgacacacacattgaagtatgcta 1784
QY 598 -----catgtccatccgcgcgag 616
DB 1785 cgagccttattcttctgtgataccttgggtacacactctcttctgatactgcgag 1844
QY 617 aagcctacatgtgtagagtggtgtctacatgatacagagccgcgtgtgagatgcctg 676
DB 1845 aatgctacttctgtcagcgtgcgtacatatacagagagcggtgagagatgcctc 1904
QY 677 aaccccccagcgctacgagattgtatataccctgtgttgatgacgaacgatac 736
DB 1905 ggtcttcctagtggtcatgagcgatgcatataccctgtgtcctgacggcgaactat 1964
QY 737 aacgagagcgactctctctccacaatgagaggtttccagctctcgtgggtgagcgt 796
DB 1965 aacgagatgtgatacctcgtctgacacgaggtgtgagacagagacactgtgggagatg 2024
QY 797 attcaagtgtagtgagccattgagatgattcagatcctagaagtatcgtatgta 856
DB 2025 atccatgtc----- 2033
QY 857 aattgtcatgtctaaccagtgctacacagacggtcagccttggcctatgctaag 916
DB 2034 -----aaggaacagcatgtgccttccctaag 2061
QY 917 tgaagcgcgcaagtaacgcttccgcttccaaagctgcgcgtcgaagctcttgcctc 976
DB 2062 tcaagcccccgaagtaacgcttccgattcctcaagcgtgcgtgtcgtgtgagctc 2121
QY 977 tgtatctgtacatcctcagagattcagagacagactcctccctcaggtatggcgtg 1036
DB 2122 tctactctgtcagagacagctctcccaagctcagaatcttcttccaagattggcctctg 2181
QY 1037 acggtgtgtcgttgaagcgctgtgtgacactgacactgtatcatctatgacgag 1096
DB 2182 atgtgtgtccttcaagcccgctcagaccttcaacttactgctgtgtgacgagc 2241
QY 1097 gctggagaggtgtta----- 1111
DB 2242 gtaacagatcattatgtatgacctccctccctcagcagaatgagtaagaacttaagac 2301

```

```

QY 1112 -----tcgaattccactctcgtgacagtcacatcgatatccgaactctctg 1162
DB 2302 taacactgttagacttccacaaacttgcgtgccagactcttgaccctgcgaactgtgtc 2361
QY 1163 gtgtcagcgtctcgtgtgtgtgagcgtgattgtgataacactgcaagatgcatgctg 1222
DB 2362 agaaccaagatgtcgcgcagcagatgagtagcgtcgcactctcgcaggtgtagtgcctg 2421
QY 1223 tctgtatgaagtcctgtagtcgcgcgaactctcaggtgtgctgtccaaactccagatg 1282
DB 2422 tctgtacgtcttgcacgtgtgag--gacacagccaggtccctccactctcgtgtgacg 2478
QY 1283 ttccttcccgaggggggggaactgtgagcccgcaaacccactgtagacagacttca 1342
DB 2479 ttccttccct-----cctacaaggaaggcccgccgcgaagaacttca 2523
QY 1343 ccttcgacgtgtcattatgacagatgacaaatcaacgagatctactctcgtatgcgaga 1402
DB 2524 agttgagcagcagcaacgagacactactgataacgattgtgtccttgcgcatgtcaatg 2583
QY 1403 accgtctgtccgcgaatgtgcccgagacactgttgagatctvgcgaactgagaacaact 1462
DB 2584 agcgtgtctcgtgcgaagcccgagctcgcgcagctgtgaggtctgagatctcgaagaactc 2643
QY 1463 ccaaggttgaactcaacccctgtcaacatcaaccctgtgtgactccgagctcttctcgt 1522
DB 2644 ctgagagctggaagcaccgcgcgtccacatcaacttgttgcattcaagaatccaaagcga 2703
QY 1523 ccaatgcgcgttgg--agtcgagcctatgagcgtgctgtgtcctcaagatgtgtctg 1579
DB 2704 ctgtgtgtctgtgcaggtcatgtccctacagatcgtgtgtcttaagatgtctgctgtg 2763
QY 1580 tggctcgtctgtgaggtgtgtatgtgtgagccacatacgtctcttccgtaagttctg 1639
DB 2764 tgggcaggggtgagaccctgtacatcagagccactcaacactctgac----- 2812
QY 1640 ccttttaccatcgttcttcaatcattatcattacaaatgaagtgtgtctacatgttga 1699
DB 2813 -----tggagcttaacatgttgca 2830
QY 1700 ctgcacaacatgatacagagagacaaacgaatgattgtgtcttcaatgtactgtct 1759
DB 2831 ctgtcaaacatcattatcagagagatacgaatgacatgtatgtcgtattcaagctaacgc 2890
QY 1760 cgtgtgataatgtgtacaaacacacgcggttcatatgacccaatgagagctctcgtgag 1819
DB 2891 ggaaggaagagatatcttcaagagagactcgaagaccatgaacccaagtgtgcgcgc 2950
QY 1820 cgcgcctctcctcgtgagagatcgaagatgtgctcgtggttgaactcagcagactgtgcat 1879
DB 2951 cgttcttcaaacacgcgaacgaacttccatgtcgtcgtcgtggaacttctccgcgcgagatc 3010
QY 1880 cactgacgcattcaagagatgtgtagcttcaaccctcaacgcagcgatgata 1933
DB 3011 cactgcccagatgcaagagctgtgcgcgagcagcgatcaaacgcctcgatga 3064

```

#### RESULT 14

AA227602 standard; DNA: 3677 BP.

AA227602:

16-DEC-1999 (first entry)

*Stachybotrys* phenol oxidase genomic sequence.

Phenol oxidase; enzyme; coloured compound; dye transfer prevention;  
fabric washing; stain bleaching; anti-dye transfer; detergent; ss.

*Stachybotrys chartarum*.



QY 1880 cactgacgcattcagagatgcttagctcaaccctacgcccaggtgatga 1933  
 ||||| ||| | ||||| ||||| | ||| ||| |||||  
 Db 3011 cactgcccaggtgcagagagctgcgcagcagcagcgtctacaccgctcgtatga 3064

RESULT 15

AA550018  
 ID AA550018 standard; DNA: 3677 BP.

AC AA550018;

DT 10-OCT-2000 (first entry)

DE Stachybotrys chartarum phenol oxidising enzyme gene.

XX phenol oxidising enzyme; detergent; bleaching; ds.

XX Stachybotrys chartarum.

XX Key Location/Qualifiers

FT CDS 1044..3095

FT /\*tag= a

FT /note= "contains introns"

FT exon 1044..1237

FT /\*tag= b

FT /number= 1

FT Intron 1238..1285

FT /\*tag= c

FT exon 1286..1324

FT /\*tag= d

FT /number= 2

FT Intron 1325..1376

FT /\*tag= e

FT exon 1377..1449

FT /\*tag= f

FT /number= 3

FT Intron 1450..1503

FT /\*tag= g

FT exon 1504..1776

FT /\*tag= h

FT /number= 4

FT Intron 1777..1835

FT /\*tag= i

FT exon 1836..3095

FT /\*tag= j

XX WO200039306-A2.

XX 06-JUL-2000.

XX 20-DEC-1999; 99WO-EP10287.

XX 23-DEC-1998; 98US-0220871.

XX 23-JUN-1999; 99US-0338723.

XX (UNIL.) UNILEVER NV.

XX (UNIL.) UNILEVER PLC.

XX (HIND-) HINDUSTAN LEVER LTD.

XX Bodie EA, Van Der Velden S, De Vries CH, Wang H;

XX WPI: 2000-514528/46.

XX P-PSDB; AAY95537.

XX Detergent composition comprising novel phenol oxidising enzyme obtained

XX from fungus or bacteria, useful for pulp and paper bleaching, bleaching

XX color of stains on fabric and for anti-dye redeposition

XX Claim 1; Fig 1A-B; 45pp; English.

XX The present sequence is that of the Stachybotrys chartarum MUC1 38898

XX phenol oxidising enzyme gene, including promoter and terminator

CC sequences. The gene was isolated from genomic DNA using primers  
 CC (see AA550023-24) based on isolated peptides of the enzyme. The  
 CC gene codes for a 594-amino acid protein (see AAY95537). The invention  
 CC relates to detergent compositions comprising novel phenol oxidising  
 CC enzymes that are encoded by nucleic acids capable of hybridising to  
 CC the present DNA sequence, provided the enzymes are capable of  
 CC modifying the colour associated with dyes or coloured compounds, and  
 CC are produced from a bacterium, yeast or fungus (see AAY95538-40). The  
 CC phenol oxidising enzymes can be used for pulp and paper bleaching,  
 CC for bleaching the colour of stains on fabric and for anti-dye  
 CC transfer in detergent and textile applications. They may also be  
 CC capable of modifying the colour in the absence or presence of an  
 CC enhancer. Expression vectors and host cells comprising a nucleic  
 CC acid encoding a phenol oxidising enzyme, methods for producing the  
 CC phenol oxidising enzyme, and methods for constructing expression  
 CC hosts are provided.

CC Sequence 3677 BP; 822 A; 1057 C; 849 G; 948 T; 1 other;

CC Query Match 15.3%; Score 300.4; DB 21; Length 3677;

CC Best Local Similarity 55.2%; Pred. No. 3.2e-72; Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

QY 316 ctctccaccagatctaccctctgagccgagccacatggttgatagcagatgcat 375  
 || || ||||| ||||| || ||||| || ||||| ||||| ||||| |||||  
 Db 1491 ctgacccttctagatctaccctctgagccgagccacatggttgatagcagatgcat 1550  
 QY 376 gtcccaagacctacatcattcgttctcgtgacatgagatggttgcgcttcgtgaa 435  
 || || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||  
 Db 1551 gaccctgtctacttacttcaatgcttccagagacagagactgagttagtctacaa 1610  
 QY 436 cagcgagagagaaacacccctccacacagctccacacagcttcttcctcagctc 495  
 || || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||  
 Db 1611 ca-----acgcacccgtgagagactcgtctcactcgcagcgcacccacgcgtgc 1664  
 QY 496 ctltgattgtggtctgagagacactaccacgctgagatgacagatgactactacc 555  
 || ||||| ||||| || || ||||| || ||||| || ||||| ||||| ||||| |||||  
 Db 1665 ttctgattgtggtctgagagactcctcccgagagacagatgactactcttc 1724  
 QY 556 caacagagagctgcgcgcatgcttctgacatgacatgc----- 597  
 || || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||  
 Db 1725 caactacaatccgcgcgcttctgtagacatgacacgcttctcagaagtatgcta 1784  
 QY 598 -----cattccatccgcgcgag 616  
 || || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||  
 Db 1785 cgagccttattcttctgctactcttctgctaaccaatctctctgtagacgtcgag 1844  
 QY 617 aacgctacatggtctcagctggtgtctacatgataccagagccgctgagatgcctg 676  
 || ||||| ||||| || || ||||| || ||||| || ||||| ||||| ||||| |||||  
 Db 1845 aatgctacttctgctcagctgcgcgctacatatacaagcagagctgagatgcctc 1904  
 QY 677 aacctccacgagctcagcgcagatttgatataccctctgcttctgactgcaagcagatc 736  
 || || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||  
 Db 1905 gttcttctcagtgctatgagagatgctgatactcttctgactgcaagcagatgact 1964  
 QY 737 aacgcagagcagctctcttccaccaatgagaggtttccagctctcgtgagtgagct 796  
 || ||||| ||||| || || ||||| || ||||| || ||||| || ||||| || |||||  
 Db 1965 aacgcagatggtacccgcgttgcagcagaggtgagagcagagactgtggagagatgctc 2024  
 QY 797 attcaatggttaagttagcccatgagatgctcagatcatcctagatagatgatatga 856  
 || || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||  
 Db 2025 atccatgct----- 2033  
 QY 857 aattgtcatgcttcaaccagtgctctacacagaaggtcagcttgctatgctcaag 916  
 || || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||  
 Db 2034 -----aacgcagacagcatgcttctccttaag 2061  
 QY 917 tgcagccgagaaatgacgcttcccttccacagcgtcgcgttcagccttcgctc 976  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 2062 tccagcccgcaagatgacttccgattccctcaagcgtgcgcgttgcgttcctc 2121  
 QY 977 tgatcttctgactcttgcagatctcagagacagactccctcaggtcattgcgctg 1036

```
Db 2122 tctaccctgcgcagacagctctcccaacg tcaagaa ttccttcccaag tcatctgcctctg 2181
QY 1037 acggtgtctgcgttgaggcgccgttgtaacactgaacactctgtacatctcatgtgcccagc 1096
Db 2182 atgcctgtctcccttcaagcccccctcaagaccttaaccttaacctgtctgttcgcgagc 2241
QY 1097 gctggagggttgta----- 1111
Db 2242 gttacgagatcatattgtatgtccctccctctcaagaa tgaagtaagactctaagac 2301
QY 1112 -----tcgaactctccacactctgccttgccagttccatcatatccgaacctctcg 1162
Db 2302 taacactgttagaacttcaaccaacttctgtgccaagactcttgacctgcgcgaactgtctg 2361
QY 1163 gtgtgtagcaggtctcgtgtgttgagccgtgaagtttataaacttgacaagatcatgcgattcg 1222
Db 2362 agaaccacagatctcgagacgagatgagtaacgctgcacactctcgaggtgagtgacgtctcg 2421
QY 1223 tctgtatgaagtccttgagtcgcgccgacacttctgaggtccttgccgaacctcgagatg 1282
Db 2422 tctgaagctctgcgcactgttgag--gacaacagccaggtccctctccactctcgttgagc 2478
QY 1283 ttccttcccccagggcgagcaacttggagcccccgaaccccaactgatagcgagacttca 1342
Db 2479 ttccttccct-----cctcacaggaaggcccccgcgcgacacgaacttca 2523
QY 1343 ccttcgagcgtgtaatgagcaatgagtaacagagttacacttctcgagatgcgaga 1402
Db 2524 agtttgaacgcagcaacgcgaactaactctgaacgagtgttgcttctgcgaatgcaatg 2583
QY 1403 accgctgtctcgcaatgtgccccgcgacactgtgagatctgcgcgacttgagaacaact 1462
Db 2584 agcgtgtcttcgcaagcccgagctcgcaacgcttgaggtctggagctcgagaactcct 2643
QY 1463 ccaacggttgagactcaacctgttcaacttactcgttgactccgaagtccttctcgt 1522
Db 2644 ctggaggctggagccaccccgctcaacttcaacttcttgacttcaagatctccaagcga 2703
QY 1523 ccaactgcgcgtg---agtcgagcctatgagctgtgcttcaagagatgttctgctgc 1579
Db 2704 ctggtgtgtctgcgcaggtcatatgccttaacgagctgtgtcttaaggaatgctcgtgct 2763
QY 1580 tggctcgtctgtaggtgtctatgttgagggcccaactaagctccttccgtaagttctcg 1639
Db 2764 tgggcaggggtgagaccctgaacatcgaggcccaactaacaaccttgac----- 2812
QY 1640 cctttaacctactgttttcaactatgctaataatctacaagtggtctacatgttga 1699
Db 2813 -----tgagacttacaatgcga 2830
QY 1700 ctgcacaacactgatccagagagacagacatgatgtgctgtcttcaatgttct 1759
Db 2831 ctgtcaaacctcatctcacagagataacgacatgatgtgctgtatccaagctcacgcgca 2890
QY 1760 cgtgtgactatgtctacaactacacgcagatcattgaaccaccaatgagacctctctgagggcc 1819
Db 2891 ggagaggaaggaatatcttcagagagacttcgagagcccatgaaccaccaagtgcgcgc 2950
QY 1820 ccggccctctccctctcgagagatctgagaaatgtgctgggtgacttcaagacttgcaat 1879
Db 2951 cgttcccttaacaacgacaacttccatctcgcgcggaacttctccgcgagttccat 3010
QY 1880 cactgacgcgcatctgagagatgtgtagcttcaaccctacgcccaagctgatga 1936
Db 3011 cactgcccagtgtagagactggtgcgcgagcagagggccgttaaacgcgctcgatga 3064
```





/db\_xref="taxon:5499"  
/clone\_lib="In vitro expressed cDNAs"  
/tissue\_type="Myxcelium"  
/note="Vector: Lambda bluescript; A mixture of mycelial  
cultures grown in liquid B5 for 48 hours and transferred  
for 24 hours to media lacking carbon, nitrogen or  
supplemented with hydrogen peroxide"

BASE COUNT 149 a 174 c 174 g 140 t 27 others  
ORIGIN

Query Match 3.5%; Score 68.6; DB 9; Length 664;  
Best Local Similarity 58.1%; Pred. No. 4.4e-06;  
Matches 158; Conservative 0; Mismatches 107; Indels 7; Gaps 2;  
QY 888 gaacggtcagccttgacatgtcgaagcgcgcgcgaatacgcctcctcct 947  
DB 57 GAACGTGCAAGCAGGCGCATATCTGAACGTGAGCCTCGCAAGTACCGATTCA 116  
QY 948 caacgtcgcgtcagccttcctcgtctgtatcttgtaacctgaagatcagagac 1007  
DB 117 TGAATGTCCTCTAGAGAGATCTTACGACATCTACATTG-----CTGATGAGAACACAA 170  
QY 1008 cgaactccctccagcagcagcagcagcagcagcagcagcagcagcagcagc 1067  
DB 171 CCCAATCCAAATTCAGGATTCGCTCAGACTCCGGCCTATTGCGCGCTTCGACGC 230  
QY 1068 tgacactctgacatctcatalgacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1127  
DB 231 CAAGACGTGGTATTTCATTCAGGAGAGCGCT-TGAGATCATTTAGACTTTTTCGTTT 289  
QY 1128 ccctgcgcagcctcagatccagcagcctc 1159  
DB 290 CCGCGGCCAAGATTTACTATGTTTCANCGGTC 321

RESULT 4  
LOCUS BE216983 848 bp mRNA linear EST 03-JUL-2000  
DEFINITION JAL5A.A06.T3 5', mRNA sequence.  
ACCESSION BE216983  
VERSION BE216983.1 GI:8904523  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Triticum.  
REFERENCE 1 (bases 1 to 848)  
AUTHORS Anderson, J.M., Williams, C.E. and Goodwin, S.B.  
TITLE Analysis of an EST database reveals a probable CF2 resistance gene  
homolog in wheat  
JOURNAL Unpublished (2000)  
COMMENT Contact: Anderson, J.M.  
Crop Production & Pest Control Research Unit  
USDA-ARS  
1150 Lilly Hall, West Lafayette, IN 47907, USA  
Tel: 765-494-5565  
Fax: 765-496-2926  
Email: janderson@purdue.edu  
Seq primer: T3  
High quality sequence stop: 848.

FEATURES  
source  
1. 848  
Location/Qualifiers  
/organism="Triticum aestivum"  
/strain="P29"  
/db\_xref="taxon:4565"  
/clone="JAL5A.A06.T3"  
/clone\_lib="Triticum aestivum Lambda Zap"  
/tissue\_type="leaf"  
/dev\_stage="9 day old seedlings"  
/dev\_stage="9 day old seedlings"

BASE COUNT 139 a 312 c 263 g 134 t

ORIGIN

Query Match 3.4%; Score 65.6; DB 9; Length 848;  
Best Local Similarity 55.2%; Pred. No. 2.7e-05;  
Matches 128; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 526 gcttcgagatagaagattactaccacacagcagcagcagcagcagcagcagcagc 585  
DB 1 GCACGAGGACTCCGCCGCGTACGATGACGACACGACAGCTCCGATACCTCGGA 60  
QY 586 ccatgacatgcacatgtccatccacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 645  
DB 61 CCACGACACGACGATGAGGCGCTCACCCGCGTCACATCCTCGGCGCTCGCGCGGA 120  
QY 646 catgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 705  
DB 121 CCGTGTCCGCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGAC 180  
QY 706 tatcccttgctcgtacgtccagcgatcacgcagcagcagcagcagcagcagcagc 757  
DB 181 CCGGACCTGCTCTCTTCGACCGGACCTTACGAGGAGCGAGCGGCGGCTTTC 232

RESULT 5  
LOCUS BF264805 880 bp mRNA linear EST 23-OCT-2001  
DEFINITION HV\_CEA0010619f Hordeum vulgare seedling green leaf EST library  
HVCDA0004 (Blumeria challenged) Hordeum vulgare cDNA clone  
HV\_CEA0010619f, mRNA sequence.  
ACCESSION BF264805  
VERSION BF264805.2 GI:13261741  
KEYWORDS EST.  
SOURCE barley.  
ORGANISM Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.  
REFERENCE 1 (bases 1 to 880)  
AUTHORS Wing, R., Close, T.J., Kleinbols, A., Wise, R., Wei, F., Begum, D.,  
Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi,  
D.W., Fenton, R.D., Oates, R. and Main, D.  
TITLE Development of a genetically and physically anchored EST resource  
for barley genomics: Blumeria infected incompatible (M1a1)  
seedling leaf cDNA library  
JOURNAL Unpublished (2001)  
COMMENT On Nov 17, 2000 this sequence version replaced gi:11195799.  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: twing@clemson.edu  
Total hg bases = 260  
Seq primer: AATTACCTCCTCACTAAGCG  
High quality sequence stop: 798.

FEATURES  
source  
1. 880  
Location/Qualifiers  
/organism="Hordeum vulgare"  
/cultivar="C116155 (M1a1)"  
/db\_xref="taxon:4513"  
/clone="HV\_CEA0010619f"  
/clone\_lib="Hordeum vulgare seedling green leaf EST  
library HVCDA0004 (Blumeria challenged)"  
/tissue\_type="seedling green leaf"  
/lab\_host="TJ121"  
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;  
C.I. 16155 (M1a1) plants were greenhouse grown in the R  
wise lab at Iowa State University, Ames, IA; 7 day old  
green seedlings were challenged with isolate A27 (AVR/M1a13  
of Blumeria graminis f. sp. hordei, and leaves were  
harvested 20 and 24 hr post-inoculation and snap frozen;









**COMMENT**

Other GSSS: RPCI-24-138E2.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@email.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi1.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tcdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tcdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 138 row: E column: 2  
Seq primer: SP6  
Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .286

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPC1-24-138B2"
/clone_id="RPC1-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pPARBAC1; Site_1: BamHI; Site_2: BamHI;
RPC1-24 Mouse BAC Library produced by Pieter de Jong. This
library was cloned in the pPARBAC1 cloning vector at the
BamHI sites using MhoI partially digested male C57BL/6J
DNA."

```

BASE COUNT	71 a	79 c	67 g	69 t
ORIGIN				

Query Match	2.48;	Score 46.2;	DB 12;	Length 286;
Best Local Similarity	52.38;	Pred. NO. 0.78;		
Matches 102; Conservative	0;	Mismatches 93;	Indels 0;	Gaps 0

479 tcttctctcgagctcccttggatgtgtggtgagacataccagcctggcgagtac 538

Db 36 TCCTCCTCCTAGTGGCAGGAGATGCCAAGATTAAGGACATCTCCTGGTTCCTCCCAAT 95

QY 539 aaggaattactactacccccaacagcgagctgccgcgatgtcttgytaaccatgacccatgcc 598

[illegible]

OY 599 atgtccatcaccqccqaaqacctacatgaqtcaqactatgtctacatgaatccaqac 658

— — — — —

[illegible]

xy 009 ccgycclgaggaclgccc 0/3

RESULT	12
BF258534	

LOCUS	BP	MRNA	LINEAR	EST	DATE
DEFINITION	BF258534	599 bp			22-OCT-2001
	HYSMET0015P21f	Hordeum vulgare seedling root EST library HVCNDNA0007			
	(Etiolated and unstressed)	Hordeum vulgare cDNA clone			
	HYSMET0015P21f,	mRNA sequence.			

ORGANISM	Hordeum vulgare
SOURCE	barley.
KEYWORDS	EST.
VERSION	GI:13119477
ACCESSION	BF258534
VERSION	BF258534.2
KEYWORDS	EST.

REFERENCE  
1 (bases 1 to 599)  
; Trilicaceae; Hordeum.  
; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poidea  
; Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
; Embryophyta; Trilicaceae; Hordeum.

TITLE	Development of a genetically and physically anchored EST resource for barley genomics: Morex unstressed seedling root cDNA library
JOURNAL	Unpublished (2001)
COMMENT	On Nov 16, 2000 this sequence version replaced gi:11187647.

Contact: Wing NA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: [twing@clemson.edu](mailto:twing@clemson.edu)  
Total hg bases = 451  
Seq primer: AATTACCTCACTAAGGC  
High quality sequence stop: 552.

FEATURES	Location/Qualifiers
source	1. .599

```

/organism="Hordium vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEF0015P21"
/clone_1b="Hordium vulgare seedling root EST library
HVCMDA007 (Etilated and unstressed)"
tissue_type="Seedling root"

```

BASE COUNT	96 a	215 c	206 g	82 f
------------	------	-------	-------	------

ORIGIN	Query Match
	2.48; Score 46.2; DB 10; Length 599;

Matches	105;	Conservative	0;	Mismatches	78;	Indels	2;	Gaps	
---------	------	--------------	----	------------	-----	--------	----	------	--

QY 889 aacggtcagccttgccctatgtcaacgtgcagccgcgcaagtlaccgcttcgcctctc 948

Db 77 AACGGCAAGCCCTGGCCGTTCTCCGCGTGCGGCGCCGCGCTACCGCTTCCGCATCCTC 136

949 aacgcctgccctctcacgcctcttcacgtctatatcttactacct--ctgaagaattcaagaa 1006

Db 137 AACGCCAGCAACGGCGCTTCTTCGGCTCTCGCTCTCGGCGGCTGGGGTTGTGCAC 196

[illegible]

Db 197 GTGGCTCCGACTCCGTGTACCTCGCCCGGCGGTGCCACGAGGGGTTCTGCTCGCG 256

0Y 1067 ctgac 1071



```

/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/notice=organ: kidney, Vector: pME18S-FL3, Site_1: DraIII
(CACCTGTGTG); Site_2: DraIII (CACCATGTGTG); 1st strand cDNA
[ATGAGGCCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TGTGTGGCCCTGTGTG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACCTGTGTG, 3' site CACCATGTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo,
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTTAAAGAGCTCG and 3' end
primer GGACCTCTCACCTCGACGACA."
122 a 137 c 120 g 106 t

```

Query Match	2.3%	Score 45.6;	DB 9;	Length 485;
Best Local Similarity	52.0%	Pred. NO. 1.5;		
Matches 102; Conservative	0;	Mismatches	94;	Indels 0;
Gaps	0;			

Oy	478	ctctctctcgcgaacccctttgtgttgccggaagaactcccaagctggcgagta	537
Db	265	CTTCTCTGTGTCAAGTGGCAGGAGATGCCAAGATTAAGACATCTCTGTTCTCCCA	324
Oy	538	caaggaattactactaccaccaagcagcagctgcccgatcttgtgtaacatgacctgc	597
Db	325	TGGGAGAAAGCTGAGGCCAAMCCAGACGGGATCTCAGTGGTGTGGAAATGATGACGATC	384
Oy	558	catgtccatcaacgcgcggaagaagccttacaatggttcagctgtgtgtctacatgatccaga	657
Db	385	CTTACCCCTCAACCATTAACAACGCCAACATGACGATGCCGGCATATACAAGGTGTGTT	444
Oy	658	cccgactagatgacgc	673
Db	445	CACGGCTGAGGACGC	460

RESULT	15
BB651243	
LOCUS	BB651243
DEFINITION	BB651243 562 bp mRNA linear EST 26-OCT-2001
	BB651243 RIKEN full-length enriched, 0 day neonate cerebellum Mus
	musculus cDNA clone C230070H21 5', mRNA sequence.

VERSION	BB651243.1	GI:16485497
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE  
AUTHORS  
Akakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, H.,

TITLE	RIEN Mouse ESTs (Arakawa, T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshinide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.go.jp/  
URL: <http://genome.gsc.riken.go.jp/>  
Yarninchi, P., Shibata, Y., Hayashi, N., Sugahara, Y., Shibata, K., Itoh

M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap trapped-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.*, 10 (10), 1617-1630 (2000)

wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carenini, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.*, 11 (2), 281-288 (2001)

Kondo, S., Shinnagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome*, 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

FEATURES

SOURCE

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="C230070H21"
/clone_lib="RIKEN full-length enriched, 0 day neonate
cerebellum"
/tissue_type="cerebellum"
/dev_stage="0 day neonate"
/lab_host="DH10a"
/note="Site_1: Salt; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGGAGAGAGAGATCCAGAGCCTTTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using triethanolamine-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 479.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGGAGAGATTCGCGATTAAATTAATATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified Bluescript KS(+) after bulk excision from Lambda
FLC I."

```

Query Match	2.3%	Score	45.6	DB	9	Length	562
Best Local Similarity	52.0%	Pred. No.	1.6				
Matches 102; Conservative	0	Mismatches	94	Indels	0	Gaps	0

OY 478 ctctctctctcgaagctccctcttgatggttgagccgaagaactaccagagctgagagta 537  
 Db 303 CTTCCTGTGTCAAGTGGCAGGAGATGCCAAAGATTAAGACATCTCTGGTTCTCCCCAA 362  
 OY 558 caagattactactaaccaccaagcagcgtgcgcgcactgtcttgatccatgaacatgc 597  
 Db 363 TGGGGAGAGAGCTGAGCCCAACACAGCAGCGGATCTCAAGTGGTGGATGATGACGACTC 422  
 OY 598 catgfcacatccgcgcgaagaagcctactatggctcagcgtgltgltacatgatccaga 657  
 Db 423 CTCTACCTCTACACATCTTTCACACGCCACACATGACAGATGCCGCAATATCAAGTGGGTGG 482  
 OY 658 ccgcgcgtgagatgccc 673

Mon Oct 7 11:28:38 2002

us-09-656-640a-1.rst

Page 10

Db 483 CACGCTGAGGACGGC 498

Search completed: October 5, 2002, 18:54:14  
Job time: 3504 sec

---

1	1958	100.0	1958	4	US-09-401-776-3	Sequence 1, Appl1
2	1946.2	99.4	2095	4	US-09-601-856-3	Sequence 3, Appl1
3	50	2.6	7218	1	US-08-232-463-1	Sequence 14, Appl1
4	39.4	2.0	2071	1	US-07-923-724-1	Sequence 1, Appl1
5	39.4	2.0	2071	1	US-08-609-426-1	Sequence 1, Appl1
6	39.4	2.0	2071	2	US-08-374-652-3	Sequence 1, Appl1
7	37.6	1.9	289	4	US-09-007-005-17	Sequence 3, Appl1
8	37.6	1.9	289	4	US-09-244-796-17	Sequence 17, Appl1
9	35.2	1.8	1704	4	US-08-528-199-2	Sequence 17, Appl1
10	35.2	1.8	1704	1	US-08-528-199-5	Sequence 5, Appl1
11	34.8	1.8	1637	2	US-08-966-316-10	Sequence 5, Appl1
12	34.8	1.8	2159	3	US-08-286-870-6-7	Sequence 10, Appl1
13	34.2	1.7	783	1	US-08-446-922-5	Sequence 7, Appl1
14	34.2	1.7	783	2	US-08-249-189-1	Sequence 5, Appl1
15	34.2	1.7	783	2	US-08-484-624-1	Sequence 1, Appl1
16	34.2	1.7	783	2	US-08-477-733-1	Sequence 1, Appl1
17	34.2	1.7	783	3	US-09-088-913A-1	Sequence 1, Appl1
18	34.2	1.7	783	4	US-08-769-819-1	Sequence 1, Appl1
19	34.2	1.7	783	4	US-08-770-974-1	Sequence 1, Appl1
20	34.2	1.7	783	5	PCT-US93-1003A-5	Sequence 5, Appl1
21	34.2	1.7	878	2	US-08-249-189-2	Sequence 5, Appl1
22	34.2	1.7	878	2	US-08-484-624-1	Sequence 22, Appl1
23	34.2	1.7	878	2	US-08-477-733B-22	Sequence 22, Appl1
24	34.2	1.7	878	3	US-09-088-913B-22	Sequence 22, Appl1
25	34.2	1.7	878	4	US-08-769-819-22	Sequence 22, Appl1
26	34.2	1.7	878	4	US-08-770-974-22	Sequence 22, Appl1
27	34.2	1.7	5392	2	US-08-403-852D-1	Sequence 1, Appl1

28	34.2	1.7	5392	3	US-08-510-646B-1	Sequence 1, Appl 1
29	34.2	1.7	5392	4	US-09-2231-818-1	Sequence 1, Appl 1
30	34	1.7	362	1	US-08-486-895-2	Sequence 2, Appl 1
31	34	1.7	443	1	US-08-486-895-1	Sequence 1, Appl 1
32	34	1.7	1248	4	US-09-105-537-7	Sequence 7, Appl 1
33	34	1.7	5370	3	US-09-105-537-7	Sequence 21, Appl 1
34	34	1.7	13613	4	US-09-103-537-3	Sequence 3, Appl 1
35	34	1.7	30001	1	US-08-123-468-1	Sequence 1, Appl 1
36	34	1.7	30001	2	US-08-474-933-1	Sequence 1, Appl 1
37	34	1.7	44033765	4	US-09-103-840A-2	Sequence 2, Appl 1
38	33.8	1.7	3417	2	US-08-464-402-1	Sequence 1, Appl 1
39	33.8	1.7	3417	4	US-09-054-775C-1	Sequence 1, Appl 1
40	33.2	1.7	752	4	US-08-818-112-2	Sequence 2, Appl 1
41	33.2	1.7	752	4	US-08-818-111-2	Sequence 2, Appl 1
42	33.2	1.7	752	4	US-09-056-556-2	Sequence 2, Appl 1
43	33.2	1.7	1086	1	US-08-235-670A-3	Sequence 3, Appl 1
44	33.2	1.7	1086	1	US-08-235-670A-1	Sequence 1, Appl 1
45	33	1.7	1524	1	US-08-173-508-1	Sequence 1, Appl 1

## ALIGNMENTS

```

RESULT 1
US-09-401-476-1
: Sequence 1, Application US/09401476
: Patent No. 6168936
: GENERAL INFORMATION:
: APPLICANT: Wang, Huming
: TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
: FILE REFERENCE: GC584
: CURRENT APPLICATION NUMBER: US/09/401,476
: CURRENT FILING DATE: 1999-09-22
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1958
: TYPE: DNA
: ORGANISM: Stachybotrys Chararum
US-09-401-476-1

```

Query Match	100.0%;	Score 1958;	DB 4;	Length 1958;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1958:	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0

QY	1	ggaatccatcaacatgatacgaacaaagatatggaacccgctggtctcgggactctgcgtgatac	60
Db	1	ggatccatcaacaacatgatatgaacaaagctatactcgaaacccgttggtctctcggggactctgcgtgatac	60
QY	61	ggcgagcaactctgtcgcgatgacatccgtctctgtcgcgatgacaggaacatatccctccggt	120
Db	61	ggcgagcaactctgtcgcgatgacatccgtctctgtcgcgatgacaggaacatatccctccggt	120
QY	121	ctcaccaagagagcagacgcaagctgagttccctcccttgagctctgtatcgaagatgctcttcgcg	180
Db	121	ctcaccaagagagcagacgcaagctgagttccctcccttgagctctgtatcgaagatgctcttcgcg	180
QY	181	atccctctctcgaagagcgcccaagatgaatgaatatactctatagctcagcagccaaagt	240
Db	181	atccctctctcgaagagcgcccaagatgaatgaatatactctatagctcagcagccaaagt	240
QY	241	tgcctaatcatgacagctacgcgtccccaacccaacacatcgggaagagacatctgtatacga	300
Db	241	tgcctaatcatgacagctacgcgtccccaacccaacacatcgggaagagacatctgtatacga	300
QY	301	gaatgagatlltagagccctctctcccaacagaatctaacctgatatctggagccgagccaacatggt	360
Db	301	gattgagatlltagagccctctctcccaacagaatctaacctgatatctggagccgagccaacatggt	360
QY	361	tgtgatacgaatgagcatgctcccaagacatacctatcatgcttccctcggtgacatgagatgt	420
Db	361	tgtgatacgaatgagcatgctcccaagacatacctatcatgcttccctcggtgacatgagatgt	420

[illegible][illegible]

```

RESULT 2
US-09-401-476-3
; Sequence 3, Application US/09401476
; Patent No. 6168936
; GENERAL INFORMATION:
; APPLICANT: Wang, Huaming
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
; FILE REFERENCE: GC584
; CURRENT APPLICATION NUMBER: US/09/401,476
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Stachybotrys Chararum
US-09-401-476-3

```

Query Match	99.4%	Score 1946, 2	DB 4	Length 2095
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1948	Conservative	0	Mismatches 3	Indels 0
				Gaps 0
QY	3	atccatacaatgatacgaagccaagctatcgcgaacgctggtcctgtcgtgatacgg	62	
Db	67	aacgatacaaatgatacagccaagctatcgaacgctggtcctgtcgtgatacgg	126	
OY	63	cggcagcctctcgaatgtgcacagatccgtgtgcctgtcgtatcgaagaagcagctctccgacct	122	
Db	127	cggcagcctctcgaatgtgcacagatccgtgtgcctgtcgtatcgaagaagcagctctccgacct	186	
OY	123	caccaagaagcagaacgacgcttgatgtcctcctcccttgcccttgtaagaagtgcctctgcgat	182	
Db	187	caccaagaagcagaacgacgcttgatgtcctcctcccttgcccttgtaagaagtgcctctgcgat	246	
OY	183	ccctcctcttgaaagcgccccaagttaagtataatctctatagctctgcagaagccaagcttg	242	
Db	247	ccctcctcttgaaagcgccccaagttaagtataatctctatagctctgcagaagccaagcttg	306	
OY	243	ctaatacttgacgaatcggctccccaacccaacaacatctgaaagagacaactctgtactacga	302	



```

|||||
Db 307 ctatcatctgacgtacgtccccaacccaacccactgagagagacatctgtactacagaga 966
QY 303 tgaagattgagccctctccacacagatcacctgatactgagccgagccaacatggttg 362
Db 367 tggagattgagccctctccacacagatcacctgatactgagagccgagccaacatggttg 426
QY 363 gatacatgacatgtccccaacccaacccaacccaacccaacccaacccaacccaacccaac 422
Db 427 gatacatgacatgtccccaacccaacccaacccaacccaacccaacccaacccaacccaac 486
QY 423 tccgcttcgtgaaacagcgagagaaacacccctcccaacagcgtccacattgacaggtctt 482
Db 487 tccgcttcgtgaaacagcgagagaaacacccctcccaacagcgtccacattgacaggtctt 546
QY 483 tccctgagctccctcttgatggttggtgagagacactaacccaacccctgagagatacag 542
Db 547 tccctgagctccctcttgatggttggtgagagacactaacccaacccctgagagatacag 606
QY 543 attactactaccccaacagcaagcgagcccgacatgcttgatccatgacatgacatgacat 602
Db 607 attactactaccccaacagcaagcgagcccgacatgcttgatccatgacatgacatgacat 666
QY 603 ccaataccgacgagaaacgacatacagtgagtgagtgagtgagtgagtgagtgagtgagtg 662
Db 667 ccaataccgacgagaaacgacatacagtgagtgagtgagtgagtgagtgagtgagtgagtg 726
QY 663 ctgagagatgccccgaacacccctcccaacgagcgagcgagcgagcgagcgagcgagcgag 722
Db 727 ctgagagatgccccgaacacccctcccaacgagcgagcgagcgagcgagcgagcgagcgag 786
QY 723 ctgcaagcgacatacagcaagcgagcgacatctctctccacaaatgagagaggtttccagct 782
Db 787 ctgcaagcgacatacagcaagcgagcgacatctctctccacaaatgagagaggtttccagct 846
QY 783 tctggggtgagcttattcaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 842
Db 847 tctggggtgagcttattcaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 906
QY 843 gatacatgataaatttgatgagtgatgacatgacatgacatgacatgacatgacatgacat 902
Db 907 gatacatgataaatttgatgagtgatgacatgacatgacatgacatgacatgacatgacat 966
QY 903 gactatgctcaacgctgacagcgagcaagtaacgacttccgcttcccaacgctgacgctc 962
Db 967 gactatgctcaacgctgacagcgagcaagtaacgacttccgcttcccaacgctgacgctc 1026
QY 963 aegctcttcgctctgatactgatacctctgagagatlcagagacagagacttccctcca 1022
Db 1027 aegctcttcgctctgatactgatacctctgagagatlcagagacagagacttccctcca 1086
QY 1023 ggtcattgacgagcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1082
Db 1087 ggtcattgacgagcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1146
QY 1083 ctctatgagcgagcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1142
Db 1147 ctctatgagcgagcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1206
QY 1143 cgaatacgcgaacacttccgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1202
Db 1207 cgaatacgcgaacacttccgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1266
QY 1203 tgaagaagtgatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1262
Db 1267 tgaagaagtgatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1326
QY 1263 gctcgcgaacacttccgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1322
Db 1327 gctcgcgaacacttccgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1386
QY 1323 cactgatgacgagacttcaacttcgagcggtgagtgagtgagtgagtgagtgagtgagtgagtgag 1382
|||||

```

```

Db 1387 cactgatgacgagacttcaacttcgagcggtgagtgagtgagtgagtgagtgagtgagtgag 1446
QY 1383 tactcttcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1442
Db 1447 tactcttcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1506
QY 1443 ctgagacttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1502
Db 1507 ctgagacttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1566
QY 1503 ctccgagcttcttcgttcacatgacgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1562
Db 1567 ctccgagcttcttcgttcacatgacgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1626
QY 1563 caagagtgatgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1622
Db 1627 caagagtgatgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1686
QY 1623 ttcccgtaagtgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1682
Db 1687 ttcccgtaagtgctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1746
QY 1683 ggtgctacatggtgacatgacgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1742
Db 1747 ggtgctacatggtgacatgacgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1806
QY 1743 ttcaatgacatggttctcgatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1802
Db 1807 ttcaatgacatggttctcgatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1866
QY 1803 gaggctctcggagggcccgccctctctctcctcgagagtgagtgagtgagtgagtgagtgagtgag 1862
Db 1867 gaggctctcggagggcccgccctctctctcctcgagagtgagtgagtgagtgagtgagtgagtgag 1926
QY 1863 tcaagcgagcttgacatgacgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1922
Db 1927 tcaagcgagcttgacatgacgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1986
QY 1923 caggctgatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1953
Db 1987 caggctgatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2017

RESULT 3
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; City: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6

```

FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)836-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE: F15  
CLONE: PT29pt-F15  
US-08-232-463-14

Query Match 2.6%; Score 50; DB 1; Length 7218;  
Best Local Similarity 0.5%; Pred. No. 0 00018;  
Matches 2; Conservative 221; Mismatches 141; Indels 0; Gaps 0;

QY 1475 ctaccctgttcacatcactcgttgactccgagtccttcgttcactgcccgtg 1534  
DB 1074 YY 1133  
QY 1535 gactgagcctatagagctgctgctcctcaagatgtgtcggcgtgctgctgag 1594  
DB 1134 YY 1193  
QY 1595 ttgtctatgttagagccgactcgtcctccctccglaagtccttcgcttacttaactg 1654  
DB 1194 YY 1253  
QY 1655 gtttctactatgtaacatctacaagtggtgtctacatgtgtcactgcacacactgat 1714  
DB 1254 YY 1313  
QY 1715 ccacgagagccacgacatgagtgctgcttcaatgctcgttcctcgagactatgagcta 1774  
DB 1314 YY 1373  
QY 1775 caactacacgagatcatgaccatgagcctctcgagagcccgcccttcctcct 1834  
DB 1374 YY 1433  
QY 1835 cggga 1838  
DB 1434 YGTA 1437

## RESULT 4

US-07-923-724-1  
Sequence 1, Application US/07923724  
Patent No. 5780292  
GENERAL INFORMATION:  
APPLICANT: Nevalainen, Helena K.M.  
APPLICANT: Paloheimo, Maria T.  
APPLICANT: Miettinen-Oinonen, Arja S.K.  
APPLICANT: Torkkeli, Tuula K.  
APPLICANT: Cantrell, Michael  
APPLICANT: Piddington, Christopher S.  
APPLICANT: Rambosek, John A.  
APPLICANT: Turunen, Marja K.  
APPLICANT: Fagerstr m, Richard B.  
TITLE OF INVENTION: Production of Phylase Degrading Enzymes  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600

CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/923,724  
FILING DATE: 31-JUL-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/496,155  
FILING DATE: 19-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/044,077  
FILING DATE: 29-APR-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 8610600  
FILING DATE: 30-APR-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimbal, Michele A.  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1050,0240004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2071 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: both  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(136..915, 970..1089, 1142..1245, 1305..1737)  
US-07-923-724-1

Query Match 2.0%; Score 39.4; DB 1; Length 2071;  
Best Local Similarity 45.6%; Pred. No. 0.12;  
Matches 139; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 391 catcatcgttctcgttgagacgtgagtggtgtcgcgttcgtggaacagcgagagacac 450  
DB 444 CAGCATCAACACTACTGATATACAGAGGCGACCTGGCTTCTGMAAGACTGGACCTACTA 503  
QY 451 ctctcccaacagcgttcacactgacagcgtctctctcgtgagctcccttgatgttgagc 510  
DB 504 CGTCCCTATGAGTGCTACTACTACAGCGCGACGACACACGCGCCCTACGCGGTTTGGT 563  
QY 511 tgaagaactaccagccctcgagcaggtacaggaattactactctaccccaagacgagctgc 570  
DB 564 GGACGGGTAAACACGACGACGATTAAGGCTGCTACAGGCGCACTCTTGAACGGTGA 623  
QY 571 ccgcatgcttggtatccatgacatgacatgcatccatccacgcgcggaacgactacaagg 630  
DB 624 GACGGTCTGCGCCCTTTCTTCTAGTGGCTACGAGCTGTCTACAGACGCGCCGCAAGTT 683  
QY 631 tcaagctggtgtcatatgataccagagcccggtgagagatgctcctgaacctccccaagcg 690  
DB 684 CGGTGAGGTTCTTGTGCTACACTACTCCACCAAGCGTCCCTCAACATCATCTCCGA 743  
QY 691 ctacg 695  
DB 744 GTCCG 748

## RESULT 5

US-08-609-426A-1  
Sequence 1, Application US/08609426A

Patent No. 5830733  
GENERAL INFORMATION:  
APPLICANT: Nevalainen, Helena K.M.  
APPLICANT: Paloheimo, Maria T.  
APPLICANT: Miettinen-Oinonen, Arja S.K.  
APPLICANT: Torkkeli, Tula K.  
APPLICANT: Cantrell, Michael  
APPLICANT: Piddington, Christopher S.  
APPLICANT: Rambosek, John A.  
APPLICANT: Turunen, Marja K.  
APPLICANT: Fagerstr m, Richard B.  
APPLICANT: Houston, Christine S.  
TITLE OF INVENTION: Production of Phytase Degrading Enzymes  
TITLE OF INVENTION: in Trichoderma  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,426A  
FILING DATE: 01-MAR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/923,724  
FILING DATE: 31-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/496,155  
FILING DATE: 19-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/044,077  
FILING DATE: 29-APR-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 8610600  
FILING DATE: 30-APR-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Reed, Grant E.  
REGISTRATION NUMBER: P-41,264  
REFERENCE/DOCKET NUMBER: 1050,0080001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2071 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: both  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(136..915, 970..1089, 1142..1245, 1305..1737)  
US-08-609-426A-1

	Query Match	Score	DB 2	Length	2071
Best Local Similarity	45.6%	Pred. No. 0.12			
Matches	139	Conservative	0	Mismatches	166
				Indels	0
				Gaps	0
QY	391	catatcgttcctcgtgcacacgagagtggtgtccgctcttgtaacacagcgagaaacac	450		
DB	444	CAGCTTCACACACTACTGTAATACAAAGGGGACCTGGCCCTTCCTGAAAGCATGGACCTTACTA	503		
QY	451	ctctcccaacagcgtccacattgcacgcgtcttctctcgcagctcccttgaatgagttgagc	510		
DB	504	CGTCCCTAATAGATGTCCTCTATCAACAGCGGAGACCAACAGCGGCCCTTCAGCGGTTTGGT	563		

OY	511	tgaggaacatccccagcctgcgcgagataaagattactatccccaaagcagagctgc	570
Db	564	ggaccgcgtacaaaccatggcmaacgattacaaagctcgcctacggccacctcttgaaacggtga	623
OY	571	ccgcatgcttgtagcaatgacacatgcatgctcatcaacgcgcgagaagcctacatg99	630
Db	624	gacggctgctgccttctttttctagtggtacgacggtcatcgaagcggcccgcaagtt	683
OY	631	tcaagctggtgtctacatgatcatcgaagcgcgcgtgagatgacctgaacctccccagc99	690
Db	684	cggtaggggtttctttggcttcaactctccacccaaagcgtgcctctcaacatcatatcttccga	743
OY	691	ctacg	695
Db	744	gtccg	748

RESULT 6  
 US-08-374-652C-3  
 : Sequence 3, Application US/08374652C  
 : Patent No. 5834286  
 : GENERAL INFORMATION:  
 : APPLICANT: NEVALAINEN, HELENA K.M.  
 : APPLICANT: PALOHAINEN, MARIA T.  
 : APPLICANT: FAGERSTROM, RICHARD B.  
 : APPLICANT: MERTTINEN-OINONEN, ARJA S.  
 : APPLICANT: TURUNEN, MARIA K.  
 : APPLICANT: RAMOSER, JOHN A.  
 : APPLICANT: PIDDINGTON, CHRISTOPHER S.  
 : APPLICANT: HOUSTON, CHRISTINE S.  
 : APPLICANT: CANTRELL, MICHAEL A.  
 : TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,  
 : TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHITATE DEGRADING  
 : ENZYMES IN DESIRED RATIOS  
 : NUMBER OF SEQUENCES: 94  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 : STREET: 1100 NEW YORK AVENUE, SUITE 600  
 : CITY: WASHINGTON  
 : STATE: DC  
 : COUNTRY: USA  
 : ZIP: 20005  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/374,652C  
 : FILING DATE: 24-MAY-1995  
 : CLASSIFICATION: 435  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US93/07058  
 : FILING DATE: 27-JUL-1993  
 : CLASSIFICATION: 435  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/925,401  
 : FILING DATE: 31-JUL-1992  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: REED, GRANT E.  
 : REGISTRATION NUMBER: 41,264  
 : REFERENCE/DOCKET NUMBER: 1050.071001  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 202-371-2600  
 : TELEFAX: 202-371-2540  
 : INFORMATION FOR SEQ ID NO: 3:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2071 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: both  
 : TOPOLOGY: linear



Sequence 2 Application US/08528199  
 Patent No. 5763228

GENERAL INFORMATION:  
 APPLICANT: KUHOTA, Michio  
 APPLICANT: TOSUAKI, Keiji  
 APPLICANT: SUGIMOTO, Toshiyuki  
 TITLE OF INVENTION: RECOMBINANT ENZYME FOR CONVERTING  
 TITLE OF INVENTION: MALTOSE INTO TREHALOSE  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: BROWDY AND NEIMARK  
 STREET: 419 Seventh Street, N.W., Suite 300  
 City: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20004

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/528,199  
 FILING DATE:

CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/485,126  
 FILING DATE: 14-SEP-1995  
 APPLICATION NUMBER: US 08/485,126  
 FILING DATE: 07-JUN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 156399/1994  
 FILING DATE: 16-JUN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BROWDY, Roger L.  
 REGISTRATION NUMBER: 25,618  
 REFERENCE/DOCKET NUMBER: KUHOTA=6A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-628-5197  
 TELEFAX: 202-737-3528  
 TELEX: 248633

INFORMATION FOR SEQ. ID. NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1704 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

US-08-528-199-2

```

Query Match 1.8%; Score 35.2; DB 1; Length 1704;
Best Local Similarity 50.0%; Pred. NO. 1.8;
Matches 88; Conservative 0; Mismatches 88; Indels 0; Gaps 0.

QY 1760 cgytgactatgctacaactacaccgagtlctatgaccccatgagacgtctctctgtaggccc 1819
Db 198 CGGGTACGACGCTCGCGGACTACACGGGATCTCTCCGGAGATCGGACGCTGAGAGACT 257
QY 1820 cgcgcctctctctccctcggaagatttcgaatgctccgggtgactctcaagcgaagcttggcat 1879
Db 258 CCACGCTTCTCCGACGGCGGCGACGAGACGCGGGATCTCGGGGATCATCGACTTGTGCTAT 317
QY 1880 cactgaccgcattcggagatgctctagctctcaaccctctacgcccgaagctgcatgaty 1935
Db 318 GAACCAACACGATGACGGCGACACCCGTGGTTCCAGGCTCTCCGACAGATCCCGACG 373

RESULT 10
US-08-528-199-5
; Sequence 5, Application US/08528199
; Patent No. 5763228
; GENERAL INFORMATION:

```

```

1  APPLICANT: KUBOTA, Michio
2  APPLICANT: TSUSAKI, Keiji
3  APPLICANT: SUGIMOTO, Toshiyuki
4  TITLE OF INVENTION: RECOMBINANT ENZYME FOR CONVERTING
5  TITLE OF INVENTION: MALTOSE INTO TREHALOSE
6  NUMBER OF SEQUENCES: 8
7  CORRESPONDENCE ADDRESS:
8  ADDRESSEE: BROWDY AND NETMAK
9  STREET: 419 Seventh Street, N.W., Suite 300
10 City: Washington
11 STATE: D.C.
12 COUNTRY: USA
13 ZIP: 20004
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.30
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/528,199
22 FILING DATE:
23 CLASSIFICATION: 435
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US 08/485,126
26 FILING DATE: 14-SEP-1995
27 APPLICATION NUMBER: US 08/485,126
28 FILING DATE: 07-JUN-1995
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: JP 156399/1994
31 FILING DATE: 16-JUN-1994
32 ATTORNEY/AGENT INFORMATION:
33 NAME: BROWDY, Roger L.
34 REGISTRATION NUMBER: 25,618
35 REFERENCE/DOCKET NUMBER: KUBOTA=6A
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: 202-628-5197
38 TELEFAX: 202-737-3528
39 TELEX: 248633
40
41 INFORMATION FOR SEQ ID NO: 5:
42 SEQUENCE CHARACTERISTICS:
43 LENGTH: 1704 base pairs
44 TYPE: nucleic acid
45 STRANDEDNESS: single
46 TOPOLOGY: linear
47 MOLECULE TYPE: CDNA
48 FEATURE:
49 NAME/KEY: CDS
50 LOCATION: 1..1704
51
52 IS-08-528-199-5

```

Query Match	1.8%	Score 35.2	DB 1	Length 1704
Best Local Similarity	50.0%	Pred. No. 1.8		
Matches 88	Conservative	0	Mismatches 88	Indels 0
				Gaps 0

  

QY	1760	cggtgactactgtgtacaaactacacccgagttcaatgaccacatgagacctctctgtgagggc	1819
Db	198	CGGGTACGACAGCTGGCGGACTACACCGGGATCCTCCCGGAGATCGGCGAGGTGAGACTT	257
QY	1820	ccggcccttcctcctcctcgtgagagttctgagaaatggtctgggttgacttcagcagatgtgcat	1879
Db	258	CCAGCCCTTCTCTCGACGGCGACGACGCGGGGATTCGGGGATCAACGATTCGTGCAT	317
QY	1880	cactgacgcagcttgcagagatggtctagcttcaaccctcagggccagagtgatgatg	1935
Db	318	GAACTACACGAGTACGCGCACCCCTGGTCTTCAAGGCTCCGCGACGATCCCGACG	373

  

RESULT	11
US-08-966-316-10/c	
; Sequence 10, Application	US/08966316
; Patent No.	5932445
; GENERAL INFORMATION:	



```
Patent No. 5716805
GENERAL INFORMATION:
APPLICANT: Spriggs, Melanie
APPLICANT: Srinivasan, Subhashini
TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,922
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/107,353
FILING DATE: 08-13-93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 1003-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 783 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
STRAIN: CD40-L
FEATURE:
NAME/KEY: CDS
LOCATION: 1..780
US-08-446-922-5

Query Match 1.78; Score 34.2; DB 1; Length 783;
Best Local Similarity 43.66; Pred. No. 2.3;
Matches 153; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
```

```
Db 534 AGAGCAGAGCTGACTGTAGCTAGCATATAGAGTCTCTCTTTAACCCTGAGCTG 475
Oy 1594 gtgtctatgttgagcccaactagctctctccgtatgtctgctt.1644
Db 474 TTTCCATTTCACAGCATYACCAAGTGTCTTTCATGTATATATCTTT 424

RESULT 14
US-08-249-189-1/C
Sequence 1, Application US/08249189
Patent No. 5961974
GENERAL INFORMATION:
APPLICANT: ARMITAGE, RICHARD
APPLICANT: FANSIOW, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU
TITLE OF INVENTION: NOVEL CYTOKINE
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.1
SOFTWARE: Microsoft Word for Apple, version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,189
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/783,707
FILING DATE: October 25, 1991
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2802-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 783 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: MOUSE
IMMEDIATE SOURCE:
CLONE: CD40-L
FEATURE:
NAME/KEY: CDS
LOCATION: 1..783
US-08-249-189-1
```





11/11/02

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 20:23:09 ; Search time 4997.71 seconds  
(without alignments)  
8772.237 Million cell updates/sec

Title: US-09-656-640A-3  
Perfect score: 2095  
Sequence: 1 cagctcggtctactctc.....gaataagttgtgtgcttaa 2095

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pal:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution:

# SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	2095	100.0	2095	6	ARI122997
2	2095	100.0	2095	6	AX101098
3	2095	100.0	2095	6	AX108674
4	1946.2	92.9	1958	6	ARI122996
5	1946.2	92.9	1958	6	AX101096
6	1946.2	92.9	1958	6	AX108672
7	528.6	25.2	1942	8	MYRBR
8	528.6	25.2	1959	8	E05283
9	517.4	24.7	2126	8	AMU271104
10	507.6	24.2	3241	8	MYRBD1
11	95.8	4.6	1422	6	AX195966
12	95.8	4.6	109519	6	AX195929
13	82.8	4.0	3446	6	AX204969
14	81.4	3.9	9531	1	AY033994
15	79.8	3.8	137879	8	AP003233
16	79.8	3.8	141528	8	AP002909
17	73.2	3.5	22302	1	SAPRSA
18	69.8	3.3	134673	8	AP002860
19	61.2	2.9	9413	1	LDMOFA
20	57.8	2.8	134673	8	AP002860
21	51.6	2.5	141528	8	AP002909
22	50	2.4	7218	6	166494
23	49.6	2.4	1533	6	AX121137
24	49.6	2.4	1614	6	AX065821
25	49.6	2.4	349980	6	AX127146
26	48.2	2.3	125020	9	AF429315
27	47.2	2.3	3170	10	RNNCAM14
28	46.4	2.2	15897	1	BSU51115
29	46.4	2.2	16585	1	AB007638
30	46.4	2.2	213190	1	BSOB0004
31	45.6	2.2	2544	10	BC011310
32	45.6	2.2	3574	4	BTADCYC
33	45	2.1	19835	4	AC092489
34	44.8	2.1	10732	6	E32986
35	44.6	2.1	10738	1	AE009694
36	44	2.1	2583	10	MNMCAMRI
37	44	2.1	2583	10	MNMCAMRI
38	43.6	2.1	127485	9	HUAC002039
39	42.6	2.0	125020	9	AF429315
40	42.4	2.0	1882	10	MMU28495
41	42.4	2.0	2031	10	MMRHOBIN
42	42.4	2.0	3131	10	AF177032
43	42.4	2.0	3154	10	BC006589
44	42	2.0	129473	2	AC040158
45	42	2.0	143900	2	AC092137

## ALIGNMENTS

RESULT 1  
LOCUS ARI122997 2095 bp DNA  
DEFINITION Sequence 3 from patent US 6168936.  
ACCESSION ARI122997  
VERSION ARI122997.1 GI:14107963  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2095)  
AUTHORS Wang,H.  
TITLE Phenol oxidizing enzymes  
JOURNAL Patent: US 6168936-A 3 02-JAN-2001;  
FEATURES  
source location/Qualifiers  
1..2095  
/organism="unknown"  
BASE COUNT 437 a 618 c 510 g 530 t  
ORIGIN

Query Match 100.0%; Score 2095; DB 6; Length 2095;  
Best Local Similarity 100.0%; Pred. No. 0;

	Matches	2095:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
QY	1	cagctcgctactactctcgtctctctctcttgcaacaatacctacacatcgcttcc	60							
Db	1	CAGCTCGGCTACTACTCTCGCTTCTCTTTGGCAAAATCAAAATCTACCAATCGTTC	60							
QY	61	atttcaaacgataccaacatgatcagcacaagctactcgagacgctgagctctgagcttgc	120							
Db	61	ATTTCAAAACGATACAAATGATATCGCCAAAGCTATCGGAGCGGTGGCTTGGCTTCTGT	120							
QY	121	gacgcgagcgacgtctctgtcgtatgcacagatccgcttgctggtgcatacgacagacatgcctc	180							
Db	121	GATCGCGGACACTCTGTGATGATCGAGATCCGTGTCTGTGTCATGCACAGACATGGCTTC	180							
QY	181	cggtctcccaagagcgagacgcaagctgagctctctccctcgagctctgttaagaagtgcct	240							
Db	181	CGGTCTCCACCAAGAGGACACCGACAGTGTGTCTCTCCCTGGCTTTGTAGAAATGGCTCT	240							
QY	241	gccgaatccctccctccgaaagcgcccaagtagtaagtaacttctatagctttagacagca	300							
Db	241	GCCGATCCCTCCTCTGAAAGGCGCCCAAGTAGTAAGTACTTATAGGTAGACAGACGA	300							
QY	301	acgtttgctaatcatatgcaggtacccgtcccaacccaacacgtgagagggagacatttctact	360							
Db	301	ACGTTGTATATATTGCAATGCAATACCGTCCCAACCCCAACTGGAGGAGACATTTTACT	360							
QY	361	acgagatggagattagggccctctctcccaacagatctacacctgatctcggagccgccaaca	420							
Db	361	ACGAGATGGAGATTAGGCCCTTCTCCACACAGATCTACCTGTATCTGGAGCCGGCCACA	420							
QY	421	tgtgttgatacagatgtgacatgtccccaagagccctacacatactgtctctgttgagcagaa	480							
Db	421	TGTTTGATATCATATGGCATGTCCCAAGACCTTACATCATCTGTCTCTGTGGCATGAGA	480							
QY	481	ggtgtgtccgcctctgtgaacacagcgagaaacaacctctcccaacagcgctccacttgcag	540							
Db	481	GGTGTGTCCGCTTCGTGAGACAGCGGAGAGAAACAACCTCTCCAAACAGCGTCCACTTGCAG	540							
QY	541	gctcttctctcgcagagctccctcttgatggtttggcttgaagaaactaacccagcctggcgagt	600							
Db	541	GCTCTTCTCTCTGAGAGCTCCCTTTGAGTGGTGGCTTGAGGACATCTACCCAGCCCTGGGAGAT	600							
QY	601	acaaagattactactaaccccccaacagcgagcttgcgcgcatgtcttggtaacatgacatg	660							
Db	601	ACAAGATTACTACTACCCCAACAGGCAAGCTGCCGCAATGCTTGGTACATGACCATG	660							
QY	661	ccatgtlccatcacccgcgcgagaaacgcctacatggttgcaagcttggcttactatgataccag	720							
Db	661	CCATGTCCATCACCGCCGAGAAACGCTTACATGGGTCAAGCTGTGTCTCATATCATCAGG	720							
QY	721	accggcgttagagatggccctgaacactcccaacggcgctagagcgaggtttgatctccct	780							
Db	721	ACCGCGCTAGAGATGGCTTGAACTCTCCACGCGGCTTACGGCGAGTTTGATATCCCTTGG	780							
QY	781	tctctgactgccaaagcagatacaacgcagacagcgacactctctctcaccaatgagagagttt	840							
Db	781	TCTCTGACTGCCAAAGCATACAAACGACAGCGGCACTCTCTTCCACCAATGGAGAGTTT	840							
QY	841	ccagctctcgggagtgacgttatctcaagtgtgaagttagtgagcccatgagatgctcagatc	900							
Db	841	CCAGCTTCTGGGGGTGACGTTATTCAAAGTGGTAGTGGAGCCCATTTGAGATGCTTCAGATC	900							
QY	901	ctagaagatgcatgtagtataatattgcatagctctaacagcgatgctatcaacagaagcgta	960							
Db	901	CTAGAGATATCGATGTATGAATTTGTGATGCTCTAACCAAGTGTATCACAGAACGCTCA	960							
QY	961	gacttggccatagtactaaagctgacgcgcgcgcaagtaacgcctcttcgccttccataacgctgt	1020							
Db	961	GACTTGGCCATAGCTCAAGTGCAGCGCGGCAAGTACCGCTTCGCTTCTCAACGCTGC	1020							
QY	1021	cgctcagaagctcttcgcctctgtatcttgtctactcttgagatctgaagacgcagacttcc	1080							
Db	1021	CGCTCAGAGCTCTTTCGCTCTGTATCTTGTCTACTCTGAGAGATTCAAGACGACTTCC	1080							

LOCUS	AX101098	2095 bp	DNA	linear	PAT 10-APR-2001
QY	1081	ctccaggatcatatgcgcgtgacaggtggtcgtcgttgaggagccctgtttgacactgacactct	1140		
Db	1081	CTTCCAGGTCATATGGCCGTGAGCGGAGGTCCTGCTTGAGGAGCCCTGTGTGACCTGACACTCT	1140		
QY	1141	gtacactctataggccgagcgctgggaggttggttatcgactctccacactccgctgagcca	1200		
Db	1141	GTACACTCTATAGCCCGAGCGCTGGGAGGTGTATATCAGCTTCTCCACCTTCGCTGGGSCA	1200		
QY	1201	gtccatcgatalcgcgaacactcctgtgtcgtcgtgacggtcgtggtgttgagcctgagtttga	1260		
Db	1201	GTCCATCGATATCCGCACACTCTCTGGTGCTGACAGCGTCTCGATGTGTGAGCCGTGAGTTTGA	1260		
QY	1261	taaacctgacaaagtcacatgcgcatttcgtttgatgaaatcctttgatgtgcgccagacatc	1320		
Db	1261	TAAACCTGACAAAGTCATGCGATTCGTTGATGAAGTCCCTGTGATGCCGACACTTC	1320		
QY	1321	tgaagtgtccctgcaacccctccgagatgttctccttcccgagagcgagcaacttggaccgcg	1380		
Db	1321	TGAGGTGCTGCGCACACTCCGAGATGTTTCCTTCCCGAGGCGGCACATGGAGACCCGC	1380		
QY	1381	aaaccgccatgtacgagagacttccactctggccgtgctaatgagacgttggacaatcaa	1440		
Db	1381	AAACCCCATGTATGACGAGACTTTCACCTTCGCGCGTCTATATGAGACATGACATCAA	1440		
QY	1441	cgaggttacctctcgagatgtcagagaaacgtctgtctccgcaaigtgcgcccgagaaactgt	1500		
Db	1441	CGGAGTACCTTCTCGAGATGTGAGAAACCGTGTGCTCCGCATATGTCGCCCGGACACTGT	1500		
QY	1501	tgaagatctgagcacttggagaaacactccaagtggttgagactaacctgtttacattcaact	1560		
Db	1501	TGAGATCTGGGCACATGGAGAACACTCCACAGCTTGGAGCTCACCTGTGTACATTACACT	1560		
QY	1561	cgttacctcgagatcccttctcgttccacatgcgcgttgagatgcgagccttatgagcgtgc	1620		
Db	1561	CGTTACTCTCGAGATCCTTCTCCTTCCACTGCCCCGTGAGTCAGAGCCTTATGAGGCTGC	1620		
QY	1621	tggctccaagatagtgtgtctgtgctgtgctgtgtgagtgtgtctatgttgaagcccaata	1680		
Db	1621	TGGTTCAGAGATGTTGTGTGCTGGCTGGCTGTGTGAGTGTGTTGTTGAGGCCACTGA	1680		
QY	1681	cgctccttcccgatgaattctcgtccttctaactaactggttttcaactatgttaaacatct	1740		
Db	1681	CGCTCCTTCCCGTAACTGTTCTGCGCTTACTGATGATTTCACTCATGCTTAACATCT	1740		
QY	1741	aaaagtgtgtctaacatgtgttcacatgcgacacccgcatccagagagaccagacataatg	1800		
Db	1741	ACAAAGTGTGTCTACATGTTCGACATGCCACACACTGATTCAGAGAGACACGACATGATG	1800		
QY	1801	gctgtcttcaatgtcaactgttctcggttgacatattgctlaactaacacgagttcatatgac	1860		
Db	1801	GCCTGCTTCAATGTCACCTGTTCTGCGGTACTATGCTATACACTGATGCACTGATTGAC	1860		
QY	1861	cccatggaacctctctcgagagcccgccccccttccttcctcctggagagatttgaagaatgctcg	1920		
Db	1861	CCCATGGAACCTCTCTGGAGAGCCCGCCCCCTTCTCTCTCTGGAGAGTTCAGAGATGGCTCG	1920		
QY	1921	ggtgacttaagagagcttttccatcaactagacgcgcatccagagagatgtgctagttcaacccc	1980		
Db	1921	GGTGACTTAAGAGAGCTTTCCTCATCTGACGCCGATTCAGAGAGATGGCTAGCTTCAACCCC	1980		
QY	1981	taagcccaagcttgatgataatgcgcgtcgtgagggtaaaatgatgatcgtctcgaaatgatta	2040		
Db	1981	TAGCCCAAGCTGATGATGATATGCCGTGAGGAGTAAATGTGATGATCGTCGATGATTTAA	2040		
QY	2041	tggacagcgagtatatagtctattttgggaaatccttgaataaatttggtgtggttaa	2095		
Db	2041	TGGACAGCAGTATATAGCTATTTTGAAGAAATCTTGAATTAAGTGTGGTGCCTTAA	2095		

DEFINITION Sequence 3 from Patent WO0121809.  
ACCESSION AX101098  
VERSION AX101098.1 GI:13619952  
KEYWORDS  
SOURCE Stachybotrys chartarum.  
ORGANISM Stachybotrys chartarum.  
REFERENCE 1 (bases 1 to 2095)  
AUTHORS Wang, H.  
TITLE Stachybotrys phenol oxidizing enzyme  
JOURNAL Patent: WO 0121809-A 3 29-MAR-2001;  
GENEMCOR INTERNATIONAL, INC. (US)  
FEATURES  
source  
/organism="Stachybotrys chartarum"  
/db\_xref="taxon:74722"  
BASE COUNT 437 a 618 c 510 g 530 t  
ORIGIN

Query Match 100.0%; Score 2095; DB 6; Length 2095;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2095; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagctcgctactactctcgtctctctcttgacaatacattcactcgtctccta 60  
DB 1 CAGCTCGGCTACTACTCTCGCTCTCTTGACAAATCTACCAATCTGCTTCA 60  
QY 61 attcaacgatacaatgatacgaagcagcagcagcagcagcagcagcagc 120  
DB 61 ATTTCAAAGATCAACATATCAGCCAACTATCGAGCGCTGCTGCTGCTG 120  
QY 121 gcttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 180  
DB 121 GATCGGCGGAGCTCTGCTGATGCCAGATCCGTTGCTGCTGCTGCTGCT 180  
QY 181 cgtctcacaagaagagcagcagcagcagcagcagcagcagcagcagcag 240  
DB 181 CGGTCTCACCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 240  
QY 241 gccgagtcctctctcgtgaagagcagcagcagcagcagcagcagcagc 300  
DB 241 GCCGATCCCTCTCTGAAAGCGCCCAAGTAAATATCTATGCTTACAG 300  
QY 301 acgttgatcatcagtcagtcagtcagtcagtcagtcagtcagtcagtc 360  
DB 301 ACCTTGCTATATCTGATACGATCCGATCCCAACCTGAGAGGAGCAT 360  
QY 361 acgagatgagatgagcagcagcagcagcagcagcagcagcagcagcag 420  
DB 361 ACGAGATGAGATGAGGCGCTTCTCCACAGATCTATCTGAGAGCGGCA 420  
QY 421 tggtttgagatacagtcagtcagtcagtcagtcagtcagtcagtcagtc 480  
DB 421 TGGTTGAGATGAGATGAGTCCCGCAGGACCTACATCTGTTCTCGTG 480  
QY 481 ggtgtgtcagcagtcagtcagtcagtcagtcagtcagtcagtcagtc 540  
DB 481 GGTGTGTCCGCTTCTGTAAGAGGAGAGAACCTCTCCCAACAGCTTCA 540  
QY 541 gctctctctcagcagtcagtcagtcagtcagtcagtcagtcagtcagtc 600  
DB 541 GCTCTCTCTCAGAGCTCCCTTGTGATGTTGGCTGAGGACACTACCA 600  
QY 601 acaagagatcaactaactaactaactaactaactaactaactaactaact 660  
DB 601 ACAAGAGATTAACAACCAACAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
QY 661 ccatgtccatcagcagcagcagcagcagcagcagcagcagcagcagcag 720  
DB 661 CCAATGCTCAACCGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 720  
QY 721 acccgagtcagagatgagcagcagcagcagcagcagcagcagcagcag 780

DB 721 ACCGGCTGAGATGCGCTTAACCTCCCGAGCGCTAACGCGAGTTGATAT 780  
QY 781 ttctgagtcgcaagcagcagcagcagcagcagcagcagcagcagcagcag 840  
DB 781 TTCTGAGTCGCAAGCATTAACGAGAGGAGGAGGAGGAGGAGGAGGAG 840  
QY 841 ccagctctcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 900  
DB 841 CCAAGTCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
QY 901 ctagaagatcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 960  
DB 901 CTAGAAGATCAAGTATGAATTTGAGCACTCTTAACAGTCTATCAAGAG 960  
QY 961 gctctgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 1020  
DB 961 GCTTGGCTATCTCAACCTGAGCGGCGCAAGTACCGCTTCTCTCAAC 1020  
QY 1021 cgtctcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1080  
DB 1021 CGTCTACGCTCTTCTGCTCTGATCTTCTGATCTGAGGATTCAGAGCA 1080  
QY 1081 ctccagagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1140  
DB 1081 CTTCAGAGTCATTCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
QY 1141 gtaacatcctcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1200  
DB 1141 GTACATCTCTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
QY 1201 gtcacatcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1260  
DB 1201 GTCCATCGATATCCGAACCTTCTGCTGCTGAGGAGGAGGAGGAGG 1260  
QY 1261 taacatcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1320  
DB 1261 TAACATGCAAGAGTCATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
QY 1321 tgaagtcgctcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1380  
DB 1321 TGAAGTCCTCGCAACCTCGAGATGTTCTTCCCGAGGCGGAGGAG 1380  
QY 1381 aaacccacatcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1440  
DB 1381 AAACCCACTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
QY 1441 cgaagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1500  
DB 1441 CGGAGTTACCTTCTGAGATGTCAGAGAACGCTGCTCCGAAATGTC 1500  
QY 1501 tgaagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1560  
DB 1501 TGAGATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
QY 1561 cgttgatcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1620  
DB 1561 CGTTGATCTCGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
QY 1621 tggttcagaagatgagtcagtcagtcagtcagtcagtcagtcagtcagtc 1680  
DB 1621 TGGTCTCAAGAGATGTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCT 1680  
QY 1681 cgtctctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1740  
DB 1681 CGCTCTTCTCCGCTTAAGTCTCGCTTAACTTAACCTGATGCTTAAC 1740  
QY 1741 acaagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 1800  
DB 1741 ACAAGTGTGCTCAATGTTGAGCTGCAACCTATCCACAGGAGCA 1800  
QY 1801 gctgcttcaatgagtcagtcagtcagtcagtcagtcagtcagtcagtcag 1860

Db 1801 GCCTTTCAATGCTACTGTTTCGGTGAATGCTACACTACACCGAGTTTCATTGAC 1860  
Oy 1861 cccatgagacccctctctgagagcccccctctccctcctcggagagttcggagatgctcg 1920  
Db 1861 CCCATGAGAGCTCTCTGAGAGGCCCCGCCCTCTCTCGAGAGATTGGAATAGCTCG 1920  
Oy 1921 ggtgactcagcagccttgcacatcagcagcagcagcagcagcagcagcagcagcagc 1980  
Db 1921 GGTGACTTCAGCGAGCTTGGCATCTACACCGCATTCAGGAGAGAGGCTTACACCC 1980  
Oy 1981 taagccagcctgtagttagtgcgcctgagagagtaataatagatcagcagatgattta 2040  
Db 1981 TACCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
Oy 2041 tggacagcagatataagcattttaggaataactgaataagtgtgtgtgttaa 2095  
Db 2041 TGGACACAGATATAGCTATTTAGGAATATCTGAAATAGTTGTGTGCTTAA 2095  
RESULT 3  
AX108674 AX108674 2095 bp DNA linear PAT 30-Apr-2001  
LOCUS AX108674  
DEFINITION Sequence 3 from Patent WO0121748.  
ACCESSION AX108674  
VERSION AX108674.1 GI:13923901  
KEYWORDS  
SOURCE Stachybotrys chartarum.  
ORGANISM Stachybotrys chartarum.  
REFERENCE 1 (bases 1 to 2095)  
AUTHORS Konvels,D.U., Doornink,M.U., de Vries,C.H. and Wang,H.  
TITLE Deletergent compositions comprising phenol oxidizing enzymes  
JOURNAL Patent: WO 0121748-A 3 29-MAR-2001;  
FEATURES  
source 1..2095  
/organism="Stachybotrys chartarum"  
BASE COUNT 437 a 618 c 510 g 530 t  
ORIGIN  
Query Match 100.0%; Score 2095; DB 6; Length 2095;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2095; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 421 tgggtgatacagatggcattgccccagagaccatccatcattcctcgtgagcattgaga 480  
Db 421 TGGTGGATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
Oy 481 ggtgtgcctcctctgtaagacgagagagagacacccctcccaacagcgtccacttgacg 540  
Db 481 GTGTTGTCGGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
Oy 541 gctcttctctcagacccctcttgaattgttggttgagagacacacacacacacacacac 600  
Db 541 GCTCTTCTCTCGAGCTCCCTTGTATGATGATGATGATGATGATGATGATGATGATGAT 600  
Oy 601 acaaggaattactaacccccacagcagcagcagcagcagcagcagcagcagcagcagcag 660  
Db 601 ACAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660  
Oy 661 ccattgcatcaccgcgcgagagagacccatacatggttcagcgtgtgtgtctacatgctc 720  
Db 661 CCATGTCATCACCAGCGAG 720  
Oy 721 acccggtcagagatgacctgagcctcccccagcagcagcagcagcagcagcagcagcag 780  
Db 721 ACCCGGCTGAGAGATGCTCTGAACTCCCAAGGCTACGGGAGTTGATATCCCTTTGG 780  
Oy 781 ttctgactgccaagcagatacacaagcagcagcagcagcagcagcagcagcagcagcagc 840  
Db 781 TTTGACTGCCAAGCAGATACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 840  
Oy 841 ccagcctctgggtgtagcatttcaagtgtagtgaagtgagccatgagatgcttcaagtc 900  
Db 841 CCAGCTTCTGGGAGAGCTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 900  
Oy 901 ctagaagatcagatgtagaagatgtagatgtagatgtagatgtagatgtagatgtagatg 960  
Db 901 CTAGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Oy 961 gacctgagcctatgctcaacagcagcagcagcagcagcagcagcagcagcagcagcag 1020  
Db 961 GCCTTGGGCTATGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
Oy 1021 cgtctcagcctcttgcctcgtatcttgcctcagcctcagagatcagagacagacattcc 1080  
Db 1021 CGTCTACAGCTCTTTCGCTGCTATCTGTAACCTTGAAGGATTAAGAGACAGACTTCC 1080  
Oy 1081 ctccaagtcattgcccgttcagcagcagcagcagcagcagcagcagcagcagcagcag 1140  
Db 1081 CTTCCAGGCTATGCTGAG 1140  
Oy 1141 gtaactctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1200  
Db 1141 GTACATCTATGAGGCGAG 1200  
Oy 1201 gtccatcagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1260  
Db 1201 GTCCATGATATCCCAACCTTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
Oy 1261 taacactgcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
Db 1261 TAACACTGACAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
Oy 1321 tgaagtgccttcgcaacccctcagagatgctctcccccagagcagcagcagcagcagc 1380  
Db 1321 TGAAGTGCTTGCACACCTTCCAGATGTTCTTCCCGAGGCGGCAACTGGGACCCCG 1380  
Oy 1381 aaacccactgatacagagacttcaacttcgacgtgtcgtatagagcagtggaacatcaa 1440  
Db 1381 AAACCCACTGATGAGAGACTTTCACCTTCCGCGCTGTATATGACAGTGAACAATTA 1440  
Oy 1441 cggagttacactctcgtgagatgtagagacagcgtctgcctcagcaatgtagcccgagaca 1500  
Db 1441 CGGAGTTACCTTCTCGAGATGCGAGAACCGTCTGCTCCCAATGTCGCCGAGACCTGT 1500  
Oy 1501 tgaagatcaggcacttgagaaacatccaaagctgagatcaccctgttcaacatccact 1560

Db	1501	TGAGATCTTGCGCACTTGGAGAACAACTCCAAACGGTTGGACTCACCCGTTCACATTACACT	1560
OY	1561	cgttaccctccgaagccttccttccttcccaactccccgtagagcgccatagaagcgyc	1620
Db	1561	CCTTAGCTTCCAGATCCTTTCCTGTTCCACTGCCGTGGAGTGCAGCCCTTAAGAAGCTTC	1620
OY	1621	tggttccaagaatgctgtgtcgtgcgtcgtcgtcgtagtgtgtctatgtttaaggccacta	1680
Db	1621	TGGTCTCAAAGATGCTTGTCTGGCTGGCTGGCTGCTGTGAAGTGTCTATGTTGAAGGCCACTA	1680
OY	1681	cgctcccttcccgaagtgtctgcgcccttaacctaactgtttccactcatgctaaacct	1740
Db	1681	GCGTCTTTCCCCGTAAGTTCCTGCTCTTTTACTTAACGTGGTTTTCACTCATGCTAACACT	1740
OY	1741	acaagtgtgtgtctacatgttgcactgcccacaaactgatctcacaggagaccaogacatg	1800
Db	1741	ACAAGTGGTGTCTACATGTTGGATGCGCAACAACCTGATCTCAGAGAGACCACACATGATG	1800
OY	1801	gctgttttcaaatgtacacggtctctcggttactatgtctaacataaccagagttcatctac	1860
Db	1801	GCTGCTTTCAAGTACACTGTTCTCGGTACTATGGCTCAACAATPAAACCGAGTTCATTAC	1860
OY	1861	cccaatggagccctcttgtagagcgcccgccctccctccctccgagaaattcgaatatgctcg	1920
Db	1861	CCCATTGGAGCCCTCTGTGAGAGGCCCGCCCTTCTCTCTCTGGAGATGTGGAGATGGCTCG	1920
OY	1921	ggttgacttcagcgagcttgcacactgaacgcacattcaggaatgtctagcttccaccoc	1980
Db	1921	GGTGACTTCAAGGACTTGGCCATCACTGACCGGCATTTCAAGANATGGCTAAGCTTCAACCC	1980
OY	1981	taagcccaagcgtatgatgatgcgcgtcgaagagaataatgatgatgcgtcogaatgat	2040
Db	1981	TACGCCCAAGGCTGATGATGATGATCCCGCTGAGGAGTAAATATGATGATCGCTGCAATGATTTA	2040
OY	2041	tggaaagagatatagctatcttttaggaaataacttgataaagtgttgtgtcttaa	2095
Db	2041	TGGACAGAGATAATAGCTAATTTTAGGAAAATCTTGAAATAGTTGTGTGTGCTTAA	2095

LOCUS	AR122996	1958 bp	DNA	linear	PAT 16-MAY-2001
DEFINITION	Sequence 1 from patent US 6168936.				
ACCESSION	AR122996				
VERSION	AR122996.1	GI:1410762			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1958)				
TITLE	Wang, H.				
JOURNAL	Patent: US 6168936-A 1 02-JAN-2001;				
FEATURES	Location/Qualifiers				
source	1..1958				
	/organism="unknown"				
BASE COUNT	394 a 593 c 490 g				
ORIGIN		481 t			
Query Match	92.94; Score 1946.2; DB 6; Length 1958;				
Best Local Similarity	99.8%; Pred. No. 0;				
Matches 1948; Conservative	0; Mismatches 3; Indels 0; Gaps 0;				
Oy	67 aacgatacaatgatcaacgaacgaatcctgagacgcttgagccttgatgacgttgcgtgatatcg	126			
Db	3 ATCCATCAACATGATGATCAACCAAGCATGAGAACCGTGCTCTGGGCTTGCTGTCGTATCGG	62			
Oy	127 cggcagcctctgcgatgcagatccgtgttcctggtgcgatcgacaagacatgcttcgcgtct	186			
Db	63 CGGCGAGCTCTGTCGATGTCGACATCGATCCGTCGTCGTCGTGATTCGACAGACATGCGCTCGGCT	122			

QY	187	caacaagagcgacgcgaagctagtcctcccttgcgcttgacaagtgctcttcgcgat	246
Db	123	CACCAAGAGGCGACGACGAGCTGAATCTCTCCCGGCTTGTAAGAATGCTCTCCGAT	182
QY	247	ccctccctcttgaaggcgccccaagtatgaatctctataggttagcnaagccaactgtg	306
Db	183	CCCTCCTTGAAAGCGCCCAAGTATGTAATGATCTATAGCTATACAAAGCCAACGTTG	242
QY	307	ctaataclttgacagctctcccacaacccaacacttgagagacactctgtactacaga	366
Db	243	CTAATCATTTGCAATGACGTATCCCAACCCCAACACTTGAAGAGGACATCTGTACTACAGAA	302
QY	367	tggagataaggccctctccacaagatactaccctgatctggaagccgccaacatglttg	426
Db	303	TGGAAATTAAGCCCTTCTCCACCAAGATCTAACCTGATCTGAGCCGCGCAACTGTTTG	362
QY	427	gataagatggatctcccccaggacctacacatagttctctgttgacctgagagtlgttg	486
Db	363	GATATGAGTGGCATCTCCCAAGACTTACACATCTCTCTGTGGCACTGAGACTGTGG	422
QY	487	tcgcgttctgtaaacgagagagacacacctctccacaagcgctcaactctgacagcctt	546
Db	423	TCCGTTGTGTAAACAGGGSAGAAACACTCTCCAAACAGCTCACTTGACAGGGCTCTT	482
QY	547	tctctcgaagctccctttgatgtgttggtctgtaggacatacccaagccttgcgagtaacaag	606
Db	483	TCTCTCGAGCTCCCTTGTATGTGTGGGCTTAGGACACTACCCAGCTGTGGCGATCAAGG	542
QY	607	attactactacccaagaagcgagctgtcccgatagctttgttaccataagcaatgcacgt	666
Db	543	ATTACTACTACCCCAAGAGGAGGCTGCCGATGCTTTGGTATCAATGACATGTGCATGT	602
QY	667	ccatcacccgcggaagcgctacatctgtgtacaagctgtgtctacaatgacacgaagccgg	726
Db	603	CCATCACCGCGGAACAGCCTACATATGGGTGAGGCTGTGTCTACATGATCCAGAACCCGG	662
QY	727	ctgagatgcccctgaacctccccaagcggtcagcgagittgatalcccccgtgtgtctga	786
Db	663	CTGAGGATGCCCTTAACCTCCCGAGCGGCTACGCGAGTTTGATATCCCTTGTTCTGA	722
QY	787	ctgccaagcgatlaaacgagacggaactctctctccacaatgagaaggtttccagct	846
Db	723	CTGCCAAGCGATTAACAGCAGACGCGCACTCTTCTTCCACCAATGAGAGGATTTCCAGCT	782
QY	847	tcctgggtggaagctatccaaagtgtgaagttgagcccatggagaatgtcttcagaatccctaga	906
Db	783	TCTGGGGTAGCTTAATTCAGTGGTAAAGTTGAGACCCATTGAGATGCTTACAGTACTTAGAA	842
QY	907	gtatcagatgatatgaattgtgcahyctctaaacagtgctatacaagaagctcagccttg	966
Db	843	GTATCAGATGTAATGAAATTTGTGCATGTGCTTAACAGATGCTATCACAAAGGTCACACTTGG	902
QY	967	ggctaibgtcaagtlgtagcgcgcaagtaacgcttcgcgttcccttaagctgcgcgtctc	1026
Db	903	GCGTATGCTCAACGTGACGCGCGGCAAGTACCGCTTCCGCTTCTTCAACGTCGCGCTCTC	962
QY	1027	aagcctcttgcctctgtatctctgtacactcttgagatacttaagagacagaactcccttcca	1086
Db	963	ACGCTCTTCCGCTCTGTATCTTGTACTCTGTAGATTTAGAGACCAAGACTTCCCTTCCA	1022
QY	1087	ggtcatgtgcgtctgaagtggtgctgtcttgaagggccctgttgaacaactgacaactctgtacat	1146
Db	1023	GGTCAATTTGCCCTGACGAGTGGTGGCTTTGAGGCGCTTGTGACACAGACACTCTGTACAT	1082
QY	1147	ctctaibgccgagcgctgggaagtttgttatacgactctccaccttcgctgcgcgaatccat	1206
Db	1083	CTTATATGGCCAGGCGTGGAGGTGTATTCAGCTTCTCCACTTCGCTGGGCACTTCAT	1142
QY	1207	cgatatccgcaacctctctgtgtctgaagcttcgcgtgtgttgaagcctgaagtgtgataaac	1266
Db	1143	CGATATCCGCAACTTCTCTGTGTCTAGCGGTCTCGGTTGTGAACCTGAGTTGATTAACAC	1202
QY	1267	tgaaagagttcatgtgacttgcgttgtaagaaagtcctttagtccgcgcgaactctcgaagt	1326

|||||  
Db 1203 TGACAAAGCATGATGATTCCTGTTGATGAAGTCCTTGAGTCGCCCAACACTTGTGAGGT 1262  
Qy 1327 gctcgcaaacctccgagatgttccctcccgagagcgagcaactgggagaccggcaaaccc 1386  
1263 gctcgcaaacctccgagatgttccctcccgagagcgagcaactgggagaccggcaaaccc 1322  
Qy 1387 cactgagagagagacttccactccgagcggtgtaatagtgagacatggaacaaacagagat 1486  
Db 1323 CACGATGACGAGACTTTCACCTTCGGCCGGCTAATGAGACAGTGAGCAATCAACGAGAGT 1382  
Qy 1447 tacccttcggagtgccgagaaacggtctgctccgcaatgtgcccggagacactgttgagat 1506  
Db 1383 TACCTTCGAGATGTCAGAAACCGTGTGCTCCGCAATGTGCCCCGCACTGTGTGAGAT 1442  
Qy 1507 ctggcgacttbaagaaacaccccaacaggtttgagacacccctgttcaacttccactcgttga 1566  
Db 1443 CTGGCGACTTGAGAACAACTCCAAACGGTGGACACACCTTTCACATTCACATCGTTGA 1502  
Qy 1567 ctccgagctcccttcgctccactgcccgttgagagtcgagaccccttaagagctgctgtct 1626  
Db 1503 CTTCGAGATCCTTCTGCTTCACTGCCCCGGAGTGGAGCTTATGAGGCTGTGCT 1562  
Qy 1627 caaggaagtgttcgctgagctgctgctgagagtggtctatgtttgagggcccaactacgctcc 1686  
Db 1563 CAAGGATGTGTGCTGCTGCTGCTGCTGAGAGTGTCTATGTGAGGCCCACTACGCTTC 1622  
Qy 1687 ttcccgtaagcttcgctcccttccactaacgtgtttcaactacatgactaactaact 1746  
Db 1623 TTTCCCGTAATCTTCGCTTTTACTACTAGTGTTCCTTACATGCTTAACTTCAAGT 1682  
Qy 1747 ggtgtctacatgttgcactgcccacaactgaltcaagagagacacagacatgagctgtct 1806  
Db 1683 GGTGTCTACATGTTGACACTGCCACAACTGATCCAGAGGACCAAGCATATGCTCTCT 1742  
Qy 1807 ttcaatgtcaatgttctggtgagctatgtgtacactacacaggttattgagaccatg 1866  
Db 1743 TTCAATGTCACTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1802  
Qy 1867 gagctctctgagagcccgccgccccttccctccgagagagttcgaagaatgctcggtgac 1926  
Db 1803 GAGCTCTCTGAGAGCCGCCGCCCTCTCTCTCGAGAGATTGAAATGGCTGGGTGAC 1862  
Qy 1927 ttcaagagagcttgcatactacatgacccgaaltcaagagatggttagcttaaccctcagcc 1986  
Db 1863 TTCAGCGAGCTTGGCATCACTGACCGCATTCAGGAGATGGCTTCAACCCCTTACGCC 1922  
Qy 1987 cagctgtagtgaatgagcgctgagagagtaaa 2017  
Db 1923 CAGCTGATGATGATGCCGCTGAGGAGTGA 1953

RESULT 5  
AX101096 1958 bp DNA linear PAT 10-APR-2001  
LOCUS AX101096  
DEFINITION Sequence 1 from Patent WO0121809.  
ACCESSION AX101096  
VERSION AX101096.1 GI:13619951  
KEYWORDS  
SOURCE Stachybotrys chartarum.  
ORGANISM Stachybotrys chartarum.  
Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Stachybotrys.  
REFERENCE 1 (bases 1 to 1958)  
AUTHORS Wang, H.  
TITLE Stachybotrys phenol oxidizing enzyme  
JOURNAL Patent: WO 0121809-A 1 29-MAR-2001;  
GENENCOR INTERNATIONAL, INC. (US)  
FEATURES  
source Location/Qualifiers  
1..1958  
/organism="Stachybotrys chartarum"  
/db\_xref="taxon:74723"

BASE COUNT 394 a 593 c 490 g 481 t  
ORIGIN

Query Match 92.9% Score 1946.2; DB 6; Length 1958;  
Best Local Similarity 99.8% Pred. No. 0;  
Matches 1948; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 67 aagatcaacaatgatacgaacagctatccgagcgctgctcgtgagctcgtgtagctg 126  
Db 3 ATTCATCAACATGATTCAGCCCAAGCTATTCGAGCGCTGCTGGCCCTGCTGATCGG 62  
Qy 127 cggcagctctgctcgtatgccaagatccgtgtgtgtagatcgagacatgctccgctc 186  
Db 63 CGGAGCTCTGTCATGAGCGCATATCCGTTGCTGTGATCGACAGACATGCTTCGGTCT 122  
Qy 187 caccgaagagcagacgagcgtgagctcctccctgctcgttgaagaagtgcctcggcagat 246  
Db 123 CACCAAGAGCGACAGCGAGTGAATCTCCCTGCTTGTACAGAAAGGCTTGGCGAT 182  
Qy 247 cccctccttgaagcgcccaagtagtaagtaacttctaaggtctagcagagccaagcttg 306  
Db 183 CCTTCCTTGAAAGCGCCCAAGTAGTATCATTTCTATAGGCTAGAGCAACGCTTG 242  
Qy 307 ctatcaatgtagtaacgctccccaaccccaacactgagagagacatctgtactaagaga 366  
Db 243 CTAAATCATTTGACATACGTCCTCCCAACCCCAACACTGAGAGAGATCTTGTACTACGAGA 302  
Qy 367 tggagattagagcccttccccaacagatcctacactgtatcttgagagcgccgccaacatgttg 426  
Db 303 TGGAGATTAGGCCCTTCTCCCAAGATCTACCTGATCTGAGCGGCGCAACATGTTG 362  
Qy 427 gataagatgagatgtccccaagagactacacatcatcgttccctcgtgtgacactgagagtg 486  
Db 363 GATACGATGAGCATGTCCCCAGAGACTACCATCATGCTTCTCTGTGACACTGAGATGTTG 422  
Qy 487 lccgctcgttgaacagcggagagaaacactcctcccaacagcgttcaacttgaagctctc 546  
Db 423 TCCGCTTGTAAGAGCGGAGAGAACACTCTCCCAACAGGCTCCACTGACAGGCTCTT 482  
Qy 547 tctctgagctcccttgaatgtgtgtggtgtaggaacactacccagagctggcgagtagaag 606  
Db 483 TCTCTGAGCTCCCTTGTATGTTGGCTGAGAGAACATACCCAGGCTTGCGATCAAGG 542  
Qy 607 attactactacccaacagagcagcgtgcccgaatgcttctgtgaacatgacatgacatgt 666  
Db 543 ATTACTACTACCCCAACAGAGAGCTGCCCGATCTTTGTGATACATACCATGCGCATGT 602  
Qy 667 ccatcacgcgcgagagacgactacatlggttcaagctgtgtgttcaatgatacgaagaccgg 726  
Db 603 CCATCACCGCGAGAAACGCTACATGAGGTCAAGGCTGAGGTCTACATGATCCAGGACCCGG 662  
Qy 727 ctgagagatgaccttgcaccccgctacgagagagttgatactccctgtgtctga 786  
Db 663 CTGAGGATGCCCTGAACCTCCCCAGCGGCTACGGGAGTTGTATATCCCTTGTGTTGTA 722  
Qy 787 ctgccaagagatacaacgagacgagcaactctcttcccaacaaatgagaggttccagct 846  
Db 723 CTGCCAAGGATACCAACGAGAGCGGACACTCTTTCACCAACATGAGAGGTTCCAGCT 782  
Qy 847 tctggggtgaagcttatccaagtgtgaagtggagcccaatgagatgcttcaagtctctaga 906  
Db 783 TCTGGGGTGACGCTTATTCAGTGAAGTTGAGCCCATTTGAATGATGATCTTCAGTAA 842  
Qy 907 gtaatgataatgaatattgtatgtcttaaacagatgctatcaagaagagtcagactgt 966  
Db 843 GATACGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 902  
Qy 967 gctatgctcaagctgcaacgagcagcaagtaacgctcctcctcccaacgctgcgtctc 1026  
Db 903 GCTTATGCTCAACGTCGACCGCGCAAGTACCGCTTCCGCTTCTCAAGCGTCCGCTTC 962  
Qy 1027 acgctcttgcgtctgtatctgttactccttgaagatltcaagagacagacttccctcca 1086  
Db 963 ACGCTCTTCGCTGTATCTTGTACTCTGTGAGATTCAGAGACGAGACTTCCCTTCCA 1022



QY 1087 ggtcattgcccgtgacggtggtctgtctttagggccctgtgacactgacactctgtaac 1146  
|||||  
Db 1023 ggtcattgcccgtgacggtggtctgtctttagggccctgtgacactgacactctgtaac 1082  
QY 1147 cttctatgcccagcgctgaggaggtgtgtatcgaactctccactctgctgagcagtcacat 1206  
|||||  
Db 1083 ctctatgagccgagccctggagaggtgtgtatcgaactctccactctgctgagcagtcacat 1142  
QY 1207 cgaattccgcaacactctccgtggtgctgacggtctggtgtgtgtgagctgaaattgtaaac 1266  
|||||  
Db 1143 cgaattccgcaacactctccgtggtgctgacggtctggtgtgtgtgagctgaaattgtaaac 1202  
QY 1267 tgacaagatgacagcattcgtctgtatgaagaatccctgtgagctgcccagacactctgagat 1326  
|||||  
Db 1203 tgacaagatgacagcattcgtctgtatgaagaatccctgtgagctgcccagacactctgagat 1262  
QY 1327 gcttgcccaacactccgagatgtctcttcccgaggagcggaactgaggaccccgcaaaccc 1386  
|||||  
Db 1263 gcttgcccaacactccgagatgtctcttcccgaggagcggaactgaggaccccgcaaaccc 1322  
QY 1387 cactgatacagagactcttcaactctgcccgtgtctatgagacagtgagacaatacagagat 1446  
|||||  
Db 1323 cactgatacagagactcttcaactctgcccgtgtctatgagacagtgagacaatacagagat 1382  
QY 1447 tacctctcagagatgacagacactgctgctcgcgaatgtgcccgcgacactgttgaagat 1506  
|||||  
Db 1383 tacctctcagagatgacagacactgctgctcgcgaatgtgcccgcgacactgttgaagat 1442  
QY 1507 ctgagcgaacttgaagaacaactccaaaggtgtgaactaacctgttcaactaacactcgttga 1566  
|||||  
Db 1443 ctgagcgaacttgaagaacaactccaaaggtgtgaactaacctgttcaactaacactcgttga 1502  
QY 1567 ctccgaagctctctctgtctgacactgcccgtgagtgagtgagacttaagagcgcgtggtct 1626  
|||||  
Db 1503 ctccgaagctctctctgtctgacactgcccgtgagtgagtgagacttaagagcgcgtggtct 1562  
QY 1627 caaagatgtgtcgtgagcgtgacgtcgtgagtgagtgagtgagtgagtgagtgagtgagtg 1686  
|||||  
Db 1563 caaagatgtgtcgtgagcgtgacgtcgtgagtgagtgagtgagtgagtgagtgagtgagtg 1622  
QY 1687 ttcccgtaagttctcgtcgtcttcaactaacactgttcaactaacactaacactaacact 1746  
|||||  
Db 1623 ttcccgtaagttctcgtcgtcttcaactaacactgttcaactaacactaacactaacact 1682  
QY 1747 ggtgtctacatgttgcacactgcccacaactgatacagagagacagacatagtgtgtctgt 1806  
|||||  
Db 1683 ggtgtctacatgttgcacactgcccacaactgatacagagagacagacatagtgtgtctgt 1742  
QY 1807 ttcaatgtcaactgtctcgtgtaactatggtctaacactaacactaacactaacactaacact 1866  
|||||  
Db 1743 ttcaatgtcaactgtctcgtgtaactatggtctaacactaacactaacactaacactaacact 1802  
QY 1867 gaggcctctctgagagcccccgcctctcctcctcctgagagatctcgaatgtgctcgtgtgac 1926  
|||||  
Db 1803 gaggcctctctgagagcccccgcctctcctcctcctgagagatctcgaatgtgctcgtgtgac 1862  
QY 1927 ttcaagcagagctgtgcatcaactgaacgcgcatctcaagagagatgtgtagcttaaacctcctagcc 1986  
|||||  
Db 1863 ttcaagcagagctgtgcatcaactgaacgcgcatctcaagagagatgtgtagcttaaacctcctagcc 1922  
QY 1987 caagctgagatgagatgacgcgctgagagataa 2017  
|||||  
Db 1923 caagctgagatgagatgacgcgctgagagataa 1953

RESULT 6  
AX108672 1958 bp DNA linear PAT -30-APR-2001  
LOCUS AX108672  
DEFINITION Sequence 1 from Patent WO0121748.  
ACCESSION AX108672  
VERSION AX108672.1 GI:13923900  
KEYWORDS

SOURCE Stachybotrys chartarum.  
ORGANISM Stachybotrys chartarum.  
REFERENCE 1 (bases 1 to 1958)  
AUTHORS Convents,D.U., Doornink,M.O., de Vries,C.H. and Wang,H.  
TITLE Detergent compositions comprising phenol oxidizing enzymes  
JOURNAL Patent: WO 0121748-A 1 29-MAR-2001;  
FEATURES  
source 1..1958  
/organism="Stachybotrys chartarum"  
/db\_xref="taxon:74722"  
BASE COUNT 394 a 593 c 490 g 481 t  
ORIGIN  
Query Match 92.9%; Score 1946.2; DB 6; Length 1958;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1948; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 67 aacgatacaactgatacagcaagctatcggagccgtggtctggtggtctgtgatacgg 126  
|||||  
Db 3 atccatcaacatgatacagcaagctatcggagccgtggtctggtggtctgtgatacgg 62  
QY 127 cggagcctctgcatgacagatccgtgtgtgtgcatgacagacatgctccgtgtct 186  
|||||  
Db 63 cggagcctctgcatgacagatccgtgtgtgtgcatgacagacatgctccgtgtct 122  
QY 187 caaccaagagcagagcagctgagctccctcctggtcctgtgaagaatgtcctctgacat 246  
|||||  
Db 123 caaccaagagcagagcagctgagctccctcctggtcctggtgaagaatgtcctctgacat 182  
QY 247 cctctctgaaagagcccaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaag 306  
|||||  
Db 183 cctctctgaaagagcccaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaag 242  
QY 307 ctaatcattgacagtaacgtccccaacccaacacacacacacacacacacacacacacacac 366  
|||||  
Db 243 ctatcatgagcagtaacgtccccaacccaacacacacacacacacacacacacacacacac 302  
QY 367 tggagatgagccctctctccacacagatctacccgtatctgtagccgcgaacatgtgtg 426  
|||||  
Db 303 tggagatgagccctctctccacacagatctacccgtatctgtagccgcgaacatgtgtg 362  
QY 427 gatagatgagatgtcccaagac 486  
|||||  
Db 363 gatagatgagatgtcccaagac 422  
QY 487 tccgctctgtagacagcggagagaacacactctcccaacagcgtcccaactgacagcgtct 546  
|||||  
Db 423 tccgctctgtagacagcggagagaacacactctcccaacagcgtcccaactgacagcgtct 482  
QY 547 tctctcagactccctctgtatgtgtgtgagtaacactcaccagcctgtagagtaagaag 606  
|||||  
Db 483 tctctcagactccctctgtatgtgtgtgagtaacactcaccagcctgtagagtaagaag 542  
QY 607 attactacaccccaacagcagcgtgctgcccagatgcttgtgatacagtaagcagatgt 666  
|||||  
Db 543 attactacaccccaacagcagcgtgctgcccagatgcttgtgatacagtaagcagatgt 602  
QY 667 ccataccgcgcgagagacgctacatagtgatcagcgtgtgtctacatgatacgaagacgg 726  
|||||  
Db 603 ccataccgcgcgagagacgctacatagtgatcagcgtgtgtctacatgatacgaagacgg 662  
QY 727 ctgagatgacccctgaaactcccaagcgtgacagcagatgtgataccctctgtgtctga 786  
|||||  
Db 663 ctgagatgacccctgaaactcccaagcgtgacagcagatgtgataccctctgtgtctga 722  
QY 787 ctgcaagcagatgatacagcagcagcagcactcttctcccaatgagagaggtttccagct 846  
|||||  
Db 723 ctgcaagcagatgatacagcagcagcagcactcttctcccaatgagagaggtttccagct 782  
QY 847 tctggagtagcttattcaagtgtagtgagtgagccatgagatgattagatctcaga 906  
|||||

D	783	TCGCGGTGACGTTATTATTCAGGTGATAGTTGAGGCCCATTTGAGATGCTTCACAGTCTGAGAA	842
Q	907	gtatcgatgataagaatttgatctgtctcctaaccagatgatatcaagaaggttcaagcttg	966
D	843	GATATCATGATATAAATTTGTGATGCTCTTCAACCAATGCTATACAGAAAGGTTGACCTTG	902
Q	967	gacctatgctcaagtcgaacgcgcgaagtaaccgtctccgcttccctcaacagcttgcgttc	1026
D	903	GCTATGCTCAACGTCGACCGCGGAAGTACGGCTTCGCGTTCTCTCAACGGCTGCGCTGC	962
Q	1027	acgcctcttcgcctcgttacctgtacctctgtagaattcaagaagacagacttcccttcca	1086
D	963	ACGCTCTTTCGCTCTGATACCTTGTACCTCTAGAGATTACAGAGAACCAACTTCCCTTCCA	1022
Q	1087	ggtcatatgcgcctggaagtgatgtctgttgaaggccctgttgaacctgaacctctgtacat	1146
D	1023	GGTTCATTGCGCGCTGACGGGTGTCTGCTTTAGGGCCCTGTGACTACTGACTCTGTACAT	1082
Q	1147	ctctatgcccgaagcgccttgggaagttgttatacgaacttccaccttcgccttgcagctcat	1206
D	1083	CTGTATGCGCGACGCTGGGAGTGTATTCACATCTTCCACCTTCGCTGCGCCAGTCCAT	1142
Q	1207	cgatatccgaaccttcctcctggtgtcgaagttcgtgtgttgaagccttgaagtttgaataac	1266
D	1143	CGATATCCGACACTTCTGCTGTACAGCTCTCGGTGTGAGCTGTGAGCTGTGATTAACAC	1202
Q	1267	tgacaagtcatacgatctcgtctgtatgaatgccttgaatgccttgcgcgaaccttctgaat	1326
D	1203	TGACAAGTCTATGGAATGTGCTGTATGAAGCTTGAAGTGTGCGCCGACACTTCTGAGGT	1262
Q	1327	gcctgcgaacctccgaagatgttctcttcccggaaggcgcaacttggagcccgcaaaccc	1386
D	1263	GCGTCCCAACCTCCGAGATTTCTTTCCTCCCGAGGGCGGCAACTGGACCCCGCAAAACC	1322
Q	1387	cactgatacgagacatttcaaccttcgcgcgtctaaatgacagttgacaatcaacgagat	1446
D	1323	CACGTATGACGAGACTTTACACTTTCGGCGGTGTATGACAGTGGACAATCAACGAGAT	1382
Q	1447	taccttctcgatgttcgagaagacggtctgtctcgcaatgtgcgcgcgcgacactgtttgagat	1506
D	1383	TACCTTCTCGGATGTGAGAACGCTGTGCTCCCGCAATGTCGCCGCACTGCTTGAGAT	1442
Q	1507	cttgcgcgaacttgaagaacaaactcgaaggtttggaactaacctctgttcaatcaactcgttga	1566
D	1443	CTGGCGACTTGGAGAACACTTCAACGCTTGAGACTACCTGTTACATTCACTCTGTTGA	1502
Q	1567	cttccgcgaatccttctcgtcttccactctgcgcgttgaagtcgaaccttayaagctgtcgtct	1626
D	1503	CTTCCGAGTCTCTTCTCTGTTCCACTCCGCTGGAGTCGAGCCCTTATGAGGCTGCTGTCT	1522
Q	1627	caagaagatgtgtcgtggtcgtctgttcgttagatgtgtctatgttttgaagcccaataacgctc	1686
D	1563	CAGAAGATGTTGCTGCTGTGCTGCTGCTGAGGTGTCATGTTGAGGCCACTTACGCTTC	1622
Q	1687	tttccgcgtaagtcttcgcaccttacctaaactatggtttttcaactatgtctaaactctcaagt	1746
D	1623	TTTCCCGTAAGTCTCCGCTTTTACCTTAACGATGATGTTTCACTCATGCTAACACTCTCAAGT	1682
Q	1747	gggtctcaacatgttgcacgtcccaaaactcgtatccaaggaagacaagacatagtatgtct	1806
D	1683	GGTGTCTACAGTGTGCACTCTCCCAACCTGTATCTACAGAGAACACACACATGATGTGCTGT	1742
Q	1807	tctcaatctactgtctcgcgttgtaactatgatacaactacaacgagtttcatcttgaaccata	1866
D	1743	TTTCATATGCACTGTTCTCGGTGACTAATGGCTTAACAACCGAGTTTATTAACCCATG	1802
Q	1867	gaacctcttggagaccgcgcaccttctctccgcgaagtttcgagaatgtgcggtgtac	1926
D	1803	GAGCTCTCTGGAAGCCCGCCCTTCTCTCTCTCGAGAGATTTCAGATATGGCTCGGGTGAC	1862
Q	1927	ttcagaagacttgcacatcaacttgaacggaattcaagaatatgtctagacttcaaccctacgc	1986
D	1863	TTTCACGCGCTTGCATCTACTGACCGCATTCAGAGATGCTTACAGTTCACACCCCTTACGC	1922

[illegible]

RESULT 7

LOCUS	1942 bp	mRNA	linear	PLN 01-FEB-2000
DEFINITION	MYRBOR M. verrucaria mRNA for bilirubin oxidase, complete cds.			

VERSION	D12579.1	GI:456711
KEYWORDS	bilirubin oxidase.	
SOURCE	Myrothecium verrucaria (strain:MT-1) CDNA to mRNA	

REFERENCE  
AUTHORS  
Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Myrothecium.  
1 (bases 1 to 1942)  
Kokkeä S., Ando, K., Kaji, H., Inoue, T., Murao, S., Takeuchi, K. and  
Samejima, T.

TITLE	JOURNAL	DATE
Molecular cloning of the gene for bilirubin oxidase from <i>Myrothecium verrucaria</i> and its expression in yeast	J. Biol. Chem. 268 (25), 18801-18809 (1993)	03/05/94

REFERENCE	2 (bases 1 to 1942)
AUTHORS	Ando, K.

TITLE	Direct Submission
JOURNAL	Submitted (07-JUL-1992) Keiichi Ando, Amano Pharmaceutical Co., Ltd.

COMMENT On Feb 26, 1994 this sequence version replaced g1:436235.

Kelichi Ando  
Tsukuba Research Laboratories  
Tsukuba, Ibaraki 305, Japan

22 Miyukigaoka  
Tsukuba  
Ibaraki 305

Japan  
Phone: 0298-56-5026

FEATURES	Location/Qualifiers
source	1. .1942

```

/strain="MT-1"
/db_xref="taxon:5532"

```

```

CDS      66. .1784
         /EC_number="1.3.3.5"

```

```

/product="bilirubin oxidase"
/protein_id="BAA02123.1"

```

```
/translation="MFKHTLGAALSLFNSNAVQASPPEPSPATGHLERVAQIASE  
QYPMETVPLPIPVKQPRLTIVNGOELIYEEVELEKPTHOYVPLDGLSVGQGO
```

PNRQ\$ARTLWYHDHAMHTTAENAYRGQAGLYMTDPAEDALNP\$SGGEEDIPMIL\$  
KQYANGMLVTNGELN\$FEM\$EDVIHVNQ\$OPW\$K\$NVE\$R\$K\$YR\$FLDA\$V\$S\$EGL\$E\$

GGSGIGIGTDIDYDNTDKVMREPVADDTTPQDTSVSPANLRDYPFSPSTNNRQFR  
GRTGPTWINGVAFADYQNRLLANYPGVGERMELINANGWTHPIHILVEKVIS

FNATVLPDYGNATVEVDPMEELWQARR  
YAADE"

```

/product="bilirubin oxidase"
1832. 1837
polyA_signal

```

BASE COUNT	440 a	549 c	436 g	517 t
ORIGIN				

Query Match	25.2%;	Score 528.6;	DB 8;	Length 1942
Best Local Similarity	62.0%;	Pred. No. 2e-118;		







Db 1509 TGTCTCGCAGCGGGTCTCTGTACGGGTGCTACGACCCGTGTGTCTGCAGCCGCTACAGTCTG 1568  
 Oy 1620 ctgtgtctcaaggaatgttgcctgtgcctgcgtcgttgaggttgctatgtlttgagcccat 1679  
 Db 1569 CTGGTCTGAAGAGACGTCGCTACCTCGGCAAGGAGAGACGGGTCTGTCGAGGCTCACT 1628  
 Oy 1660 acgcctccttcccgtaagttctgccttacctactcgtgttlttaactaagtctaacc 1739  
 Db 1629 ACGCTCTTGTGGCC----- 1642  
 Oy 1740 tacaagtgtgtgtcatcatgttgcactgtccacaaactgtatccaagagcaacagatgat 1799  
 Db 1643 -----GGCGGTACATGTTCTCACTGGCACACTCATTCCACGGAGCAAGACATGAT 1695  
 Oy 1800 ggcctgttccaatgtcactgcttctcgttgaaactatgtctacaactacaacagatcattta 1859  
 Db 1696 GGGCGGCTTCAACAGTCACCGTCTCTCCGAGACTACGAGCTACAACTGGACGGCCCTTGCTGA 1755  
 Oy 1860 ccccaatgagcctctcttgaaaggcccgcccttccctcctcctcgtgagatctcgaaatgctc 1919  
 Db 1756 CCCCATGCAAGGACGAGTTCCGGCGCAAGCGGTACGTGGCAACGACGTCGAGGTCCGGGA 1815  
 Oy 1920 ggggtgacttcacgagagcttgcacatcaactagcacgcaatcaagagaatgctgagctcaacc 1979  
 Db 1816 GAACGCCCTTCAAGCACCGACGAGATGAGGCGCAGGTCCAGCTGATGCGCAGCTACTCTGCC 1875  
 Oy 1980 ctacgcccagagctga 1994  
 Db 1876 CTACGACAAACCGGA 1890

```

RESULT 10
MYRBD
LOCUS MYRBD 3241 bp DNA linear PLN 14-APR-2000
DEFINITION M. verrucaria DNA for bilirubin oxidase, complete cds.
ACCESSION D14081
VERSION D14081.1 GI:436236
KEYWORDS bilirubin oxidase.
SOURCE Myrothecium verrucaria (strain:MT-1) DNA.
ORGANISM Myrothecium verrucaria
Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Myrothecium.
REFERENCE 1 (bases 1 to 3241)
AUTHORS Kojkega,S., Ando,K., Kaji,H., Inoue,T., Murao,S., Takeuchi,K. and Samejima,T.
TITLE Molecular cloning of the gene for bilirubin oxidase from Myrothecium verrucaria and its expression in yeast
J. Biol. Chem. 268 (25), 18801-18809 (1993)
MEDLINE 93366794
REFERENCE 2 (bases 1 to 3241)
AUTHORS Kojkega,S.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1993) Satoshi Kojkega, Amano Pharmaceutical Co., Ltd., Tsukuba Research Laboratories, 22 Miyukigaoka, Tsukuba, Ibaragi 305, Japan (Tel:0298-56-5026, Fax:0298-56-5012)
COMMENT Submitted (20-JAN-1993) to DDBJ by:
Satoshi Kojkega
Tsukuba Research Laboratories
Amano Pharmaceutical Co., Ltd.
22 Miyukigaoka
Tsukuba, Ibaragi 305
Japan
Phone: 0298-56-5026
Fax: 0298-56-5012.
FEATURES
Source
Location/Qualifiers
1..3241
/organism="Myrothecium verrucaria"
/strain="MT-1"
/db_xref="taxon:5532"
260..265
/note="TATA box-like sequence"
706..711
TATA_signal
973..978
TATA_signal

```

	misc_feature	/note="TATA box-like sequence"
exon	1011..1291	/number=1
	1011	
	/note="major transcription initiation site"	
CDS	join(1104)..1291,1353..1504,1564..1955,2036..2608, 2677..2826,2878..3141)	
	/EC_number="1.3.3.5"	
	/codon_start=1	
	/product="Bilirubin oxidase"	
	/protein_id="BA03166.1"	
	/db_xref="gi:456710"	
	/translation="MEKTLGAALAILFNSNAVOASPPEPSPATGHLFRVADISL QYPMETVPLPIEPVKQPRLLTYINIPYNGEILWYELEIFTHQVPDGLSADLVITDD MSPGTFQVPGSEVETVRFENNAEAPNSVGHLHGFSRAAFDMADITEPOSEKDYVY PNRSARSRLWMYHDHAMHITTAENAYRGOAGLYMLTDPADDALNLPGEVEFDIPLILT KOYTANGMLVTTNELSNFMGDIVHNVCQPMFEKVFREPRKAEFRELDAAVSFGLY ADTIDIRLPERKVIASDSGLEHPADTSLEYISMAERYEVDFEDSDYAGKTIELRNLI GGSTIGETDDNDYONTDKVMRFVADVDDTPOTSVYPANLRDPVPSPTNTNRQRR GRGTGWTINGVAERADYONRLLANYPVGTVERMELINANGNTHFIHLHDFKYSIS TSGNAKRTVMPEESGLKDVVLGRETIVVEAHNAPEGVYEFCHNLIHEDDMAA FNATVLPDYGNATVEVDPMELMWAREYELDEFOASOGFSVCAVTERIQTMAYERIK	
	YAADE"	
sig_peptide	1104..1217	
mat_peptide	join(1218..1291,1353..1504,1564..1955,2036..2608, 2677..2826,2878..3138)	
	/product="Bilirubin oxidase"	
intron	1292..1352	
	/number=1	
exon	1353..1504	
	/number=2	
intron	1505..1563	
	/number=2	
exon	1564..1955	
	/number=3	
intron	1956..2035	
	/number=3	
exon	2036..2608	
	/number=4	
intron	2609..2676	
	/number=4	
exon	2677..2826	
	/number=5	
intron	2827..2877	
	/number=5	
exon	2878..>3241	
	/number=6	
POLYA_signal	3189..3194	
BASE COUNT	803 a 848 c 745 g 845 t	
ORIGIN		
Query Match	24.2%	Score 507.6; DB 8; Length 3241;
Best Local Similarity	58.7%;	Pred. NO.2.8e-113;
Matches 1218; Conservative	0;	Mismatches 669; Indels 189; Gaps 12
Dy	73 caacatgatcaaccaagcatcagagccgctgccttcggccttgctgtagtaccgagcaag	132
Dd	1100 CAAGATGTTCAACAACACACTTGAGAGCTGTGCCCTCACCATT---GCTTTTCACAGCA	1156
Oy	133 ctctgctgatgccagatccgttgtcgtgctgatcgacaagaatgacctccggtctcaacaa	192
Dd	1157 TGCCTGTCCAGCGAACACCCCGTCCCGAGACACCTCACCGGCAA--CTGGACATCTCTTCAA	1213
Oy	193 gagcagcagcagcagctgaattccctcccctggcctttgtaagcgaagtgccttcgcatccctcc	252
Dd	1214 GCGAGTTGCCCGACGATCAGCCCACTATTCCTCATGTTTCACATGATCACCAATTCCTTC	1273
Oy	253 tctgaaggcgcaccaagttagtatcatctctaagcctagcagcagcagcagcttgctcta---	309
Dd	1274 TCTTAAGCAGACCCCCCTGTGTGAGTAGACGATTTTCAAAGTGCATATCTCATATGAAAGAATAT	1333
Oy	310 -----atcatgtcacgtaccgtcccccaacccaacacactlgaagaggacatcttgcact	360

Db	1334	GCTGACTCTATGGAATGGAAGGACTGTAAACCAATCTCTGTGAATGGACAGAGATCTGGTACT	1393
OY	361	acgagatggagattagagccctctcccaagaatctaacctgatctggagcggcaaca	420
Db	1394	ATGAGGTGAGATCAAGCCCTTCACACTACACAGGATTACCTGTACCTTGATTCGGCTGATC	1453
OY	421	tggatggatagagagcgtgcccaagagactacatcatcgttccctgtagaac-----	475
Db	1454	TGGTCGGGTATGATGGAATGTCCTCTGGCCCTACTTTCCAGATTCCCTGATGTAAGCTT	1515
OY	476	-----tgagag	481
Db	1514	CGAGACCATGCAAAAAGGATTTGAATAATGAGCGCTGATTAGAAACACAGAGATTGAAAC	1573
OY	482	tgtgtccgctctgtgaacgacggagagaacacctctcccaagcgctcaattgcaag	541
Db	1574	AGTTGTCCGCTTCATTAACAAATGCTGAGG-----CTCCTAACCTCCCTTACCTGCAGCG	1627
OY	542	ctctctctcgaagctccctcttgatgtggctggaagaaactaacagcctgagat	601
Db	1628	ATCATCTCTCTGTGCGCGCTTTGACGGATGGCGAGAGACATACCGACGCTTGGACGTT	1687
OY	602	caaggaattactaaccaccaagcagcgtctcccgatgcttgbtaccatgacatgc	661
Db	1688	CAAGACATATTACTACCCAAATAACACAGCTGTCTGTACCTGTATGGATACCAACATATGC	1747
OY	662	catgtccatcaccccggaagacgtctacaatgbgtccagctggtgtctacatgttccaagg	721
Db	1748	TATGATATTCTACTGCTGAGAAACGCTTACCGTGGCCAGAGCTGTGTTCTTACATGCTCACTGA	1807
OY	722	cccgcttgaagatgccttgaacctcccccagcgctgaagcgagattgatalcccttgt	781
Db	1808	CCGAGCCCAAGACGCTCTCAACTGTGCAAGAGGTATGGGAGATTCATATTCCATATGAT	1867
OY	782	tctgactgcgaagagataaacagagagcgacactctcttccaccaatgaaaggttcc	841
Db	1868	CCTCAGCTCAAGCAATATTACCGCAAAAGGCAACTTGTGTACCACTATATGAGAAAGCTGAA	1927
OY	842	cagcttcggggatgacgttattcaagtgtgaagtgtgagcccatlgagatgcttcaatcc	901
Db	1928	CTCATTTCTGGGGTGAATGTAATTCAGTGTGATACGTAATGAAGCAATCTGTATGACCGGT	1987
OY	902	tagaagtatcgatgtatgaaattgtgtcatgtcttaaccagttgctatcacagaacggtcag	961
Db	1988	ATACAAATTCAACCCCCCATTTAAACAGTACGACATATCGAACAACGTGT---AGAAACGGTCAA	2044
OY	962	ccttgagcctatgctcaacagtcagccggcgcaagatcacgccttccgcttccccaagcgtgcc	1021
Db	2045	CCCTGCGCTTTCAAGAACGTTGAGACCTTCGCAAAATATGTGATTTCCGCTTCTCGATATGCCGA	2104
OY	1022	gtctcaagctcttctgcgtctgtatcttgtctaacctctgaggaattcaagacaagactccc	1081
Db	2105	GTTTCTCTGCTCTTTCGGGCTTTACTTTGTGTGATACTGATCTATACGACATCGCTTGCT	2164
OY	1082	ttcgaagttatggccgtcgaacggtggtcgtcgtcttgagggccctgttgacactgacactcg	1141
Db	2165	TTCAAGGTTATTGGCTCGGATTTCTGGTCTCTTGAACAACCTTCGCCGATPACACTTGTCTG	2224
OY	1142	tacatctctatgagccagcgtctgggaagttgttlatcgaactctccacactctgcctggcag	1201
Db	2225	TACATTTTCATGAGCGCGAGCGTTACGAGATTGTGTGATCACTTTCGCCATATTGCTGGGAAAG	2284
OY	1202	tccaatgatataccgcaactctccgtgt---gctgaagctctggatgtgttgagcctgagttc	1258
Db	2285	ACTATTGATGACTCCGAACCTTGGCGGTGAGCATTTGGGCGCATTCGGAAACAGATTACGACTAT	2344
OY	1259	gataaacatgacaaggtcatgcatctgctcgttataaagtctccttgaagtcggccgaact	1318
Db	2345	GACAACACCGAACAGGTATGCTGTTGTGTGTACACACACACAACTCAGCCAGATATAC	2404
OY	1319	tctgagtgctctccaaccttcgagatgttctcttcccgaaggcggcaacttggagcccc	1378

LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM
DB 2405	TCAGTTGTTCCGTCAACCTTGATGTTGCTTCCCTTCCCTTCGCCACCAACAACACCC	1438			
OY 1379	gcaaaccccaactgatatgacagagacttccactctcggcgtgtcaatgagacagtgacaac	1438			
DB 2465	CGACA-----GTTCCGGCTTTGGTTCGGACCGCGATCCATACCGTCACTAT	2506			
OY 1439	aacggagttacactcttcgtagtgcgagacccgtcgcctccgaatgtgtcccgagact	1498			
DB 2507	AATGGTGTGCTTTTGGCTGATGTTTCAAAACCGCTGTGCTTGCACAACGTACCGTGTACT	2566			
OY 1499	gttgagatctggcgcactgtgagacaactcccaaggtgtgac-----	1539			
DB 2567	GTCGAGCGCTGGAGCATCATCAACCGCGTACGGTTGACGCGTATGTTTCTTCGGAAT	2626			
OY 1540	-----tcacccctgttc	1550			
DB 2627	ATGACTACCCCTTGCTAGATAATTTGAGAGAATTTGCACATAACACTTCAGACCCCTATTC	2686			
OY 1551	acaattcacctcgttgaactctcgaatcccttccttcgttc-----caatgcgcgtgag	1601			
DB 2687	ACATCCATCTTGTCGACCTTCAAGGTGATTTCTCGTACTTCGCGCAACAACGGCGCACAG	2746			
OY 1602	tcgagccctatgagagcgtcgtgtctcaagatgltgtcttgactgtgcgtcgtgtaggtg	1661			
DB 2747	TCATCCCAATACAG--TCGCGTCTCAAAAGAGAGTGTGTGCTGTGGTGGCGCTGAAACTG	2803			
OY 1662	tctaattgagagcccaactagcctcccttcocgtaagcttctgcgttccttttaacctga	1721			
DB 2804	TGGTGTGTTGAGGCTCATTTACCGCGGTATGACTGTTATTCATGCCAAATTTACAGCTGAT	2863			
OY 1722	ttcactcatgtcctaactatcacaaagtggtgtctacatgttgacatgacgcacaactgatca	1781			
DB 2864	CTAATATGCGGGTAGCGCTTTCC--TGCGTATACAGTTGTTCCATTCGCAATTTGATCA	2921			
OY 1782	cgagagaccagacatgatatgctgccttcaatgtacactgttcccgtagtactatgataca	1841			
DB 2922	CGAGGATACGATATGATGAGGCTGCTTTACGCCACCGCTCGCCAGATATAGGCTATTA	2981			
OY 1842	ctaacccagatctcatctgaccccatatgagacccctctcgaagagcccgccccctctcc	1901			
DB 2982	TGCCACTGTTTTGCTTGAACCTATGGAAGAGCTTTGGCAGCGCTGTCTTATGAACCTGG	3041			
OY 1902	agagttcgagaatgctgcgtgtgacttcaagcagagcttgccatacactgacgcgcatcaga	1961			
DB 3042	CGAGTTCCAGGCTCGAGATGGCCACTTACGCGCTGTACTGAGCGATATTCAGAC	3101			
OY 1962	gatgctagcttcaaccctacgcccagcgtatga 1997				
DB 3102	TATGCTGATATACAGACTTACGCCGCGACGCTACGA 3137				
RESULT 11					
AX195966	1422 bp	DNA	Linear	PAT 28-AUG-2001	
LOCUS	AX195966				
DEFINITION	Sequence 38 from Patent WO0151639.				
ACCESSION	AX195966				
VERSION	AX195966.1				
KEYWORDS	GI:15386198				
SOURCE					
ORGANISM					
	Micromonospora carbonacea.				
	Micromonospora carbonacea				
	Bacteria: Firmicutes; Actinobacteria: Actinobacteridae;				
	Actinomycetales; Micromonosporinae; Micromonosporaceae;				
	Micromonospora.				
	1 (bases 1 to 1422)				
REFERENCE					
AUTHORS	Hosted,T.J., Horan,A.C. and Wang,T.X.				
TITLE	Evernimycin biosynthetic genes				
JOURNAL	Patent: WO 0151639-A 38-19-JUL-2001;				
	Schering Corporation (US)				
FEATURES					
SOURCE					
	1..1422				
	/organism="Micromonospora carbonacea"				
	/db_xref="taxon:47853"				
	1..>1422				

```
/note="evic"
/codon_start=1
/transl_table=1
/protein_id="CAC6009.1"
/db_xref="GI:15386199"
/translation="MSRSLRDAQAQAAPASPNPHNGHAPVPSRVSTTVAVTP
TEPVPVPRRLTPVSRDIDIVEIPIRAQVQIILPEGLTPATVAGSVGPTIRATW
RPRVITTYNGLDTHANVHLHGHWAPATSDHPMDLIPGSGKVDYDPMLOGATLWHT
DHTHAYEADHYRGLHGFYLLIDDEPAEHLRLRPAKGYDVPIMLRNAQEDSGALVEGHP
DDRVTILANGKAQYFEVAPRRYRFRHLNALKHYFRLNLGSEPLRTATDGLLPAP
TSHNELALSSGEREIVYIDFAHAGSGPVLYDGNPLTRFDVSSRATDSDSRVPLT
RALPDMGTPTVETRVSMSEFDSARPRLIMGMKFPDPLRVQVKKRSTELMNVNAD
TDPEPDPHFLHLVTRFVRLGRDGPAPPEADGLADIVYVSPKSKIQVTFATPYLGS
QYVYHCHYLEHSSLSGMAQLEVP"

BASE COUNT      219 a      575 c      431 g      197 t
ORIGIN

Query Match      4.6%; Score 95.8; DB 6; Length 1422;
Best Local Similarity 54.8%; Pred. No. 1.3e-12;
Matches 212; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

Oy 472 gcaatgagagtggttcgcgttcgtgaacagagagagagacacccctcccaacagcgtcc 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 GCACGGGCGGCCGCTGGCGATCCTACACCAACGGGCTCGACACCCGCAACGTC 358

Oy 532 acttgcagagctcttcctcagctcccttgatgttggtgagagagacacacagc 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 359 ACCTGACAGCGCGGACGCGCGGACACAGGAGGACGCGCATGACCTGATCCCGC 418

Oy 552 ctggcagatcaaggattactactaccocaaagcagagcgtcccgcatgtcttggtacc 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 419 CGGCGGCTCGAAGGCTACAGCACTACCGAACCTTCAGGCGGCGCGAGCTCTGTGTAAC 478

Oy 652 atgaccatgcacgtccatcacccgcagagacgtctacatggtgaggtgtgtactata 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 479 ACGACCAACACCCAGCCTTACAGCGCGACGACGCTACCGCGGACTGACAGGCTTCTATC 538

Oy 712 tgatcagaagaccgcgtgagatgctccctgaaccccccagcagcgtacagcagattgata 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 539 TGATCGACAGACCGCGCGGACGATCAGCTGCGCTCGCGCGGCAAGTACAGACGTCGCA 598

Oy 772 tcccttggtcttgactgacacagcatatacagcagagcagcactctctctccaccatg 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 599 T---CATGCTGGCAACGCCGCCGTCGACGACTCCGCGCCCTCGCTTGGCGCACCCGCG 655

Oy 832 gagaggttccagctctcggtggtgacg 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656 ACGACCGGCTCACATCTCTGGCGAAGC 682

RESULT 12
AX195929      109519 bp      DNA      linear      PAT 28-AUG-2001
LOCUS      AX195929
DEFINITION      Sequence 1 from Patent WO0151639.
ACCESSION      AX195929
VERSION      AX195929.1 GI:15386161
KEYWORDS
SOURCE      Micromonospora carbonacea.
ORGANISM      Micromonospora carbonacea.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Micromonosporinae; Micromonosporaceae;
Micromonospora.
REFERENCE      1 (bases 1 to 109519)
AUTHORS      Hosted,T.J., Horan,A.C. and Wang,T.X.
TITLE      Evernimicin biosynthetic genes
JOURNAL      Patent: WO 0151639-A 1 19-JUL-2001;
Schering Corporation (US)
FEATURES
source      1..109519
              /organism="Micromonospora carbonacea"
              /db_xref="taxon:47853"
BASE COUNT      15101 a      39910 c      39406 g      15102 t
```

```
ORIGIN

Query Match      4.6%; Score 95.8; DB 6; Length 109519;
Best Local Similarity 54.8%; Pred. No. 1.6e-12;
Matches 212; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

Oy 472 gcaatgagagtggttcgcgttcgtgaacagagagagagacacccctcccaacagcgtcc 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23046 GCACGGGCGGCCGCTGGCGATCCTACACCAACGGGCTCGACACCCGCAACGTC 23105

Oy 532 acttgcagagctcttcctcagctcccttgatgttggtgagagagacacacagc 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23106 ACCTGACAGCGCGGACGCGCGGACACAGGAGGCTACCCGATGACCTGATCCCGC 23165

Oy 592 ctggcagagatcaaggattactactaccocaaagcagagcgtcccgcatgtcttggtacc 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23166 CGGCGGCTCGAAGGCTACAGCACTACCGAACCTTCAGGCGGCGGACGCTCTGTGTAAC 23225

Oy 652 atgaccatgcacgtccatcacccgcagagacgtctacatggtgaggtgtgtactata 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23226 ACGACCAACACCCAGCCTTACAGCGCGACGCTACCGCGGAGTACGACGCTTCTATC 23285

Oy 712 tgatcagaagaccgcgtgagatgctccctgaaccccccagcagcgtacagcagattgata 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23286 TGATCGACAGCGCGGACGCGGACGATCAGCTGCGCTCGCGCGCAAGTACGACGTCGCA 23345

Oy 772 tcccttggtcttgactgacacagcatatacagcagagcagcactctctctccaccatg 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23346 T---CATGCTGGCAACGCCGCCGTCGACGACTCCGCGCCCTCGCTTGGCGCACCCGCG 23402

Oy 832 gagaggttccagctctcggtggtgacg 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23403 ACGACCGGCTCACATCTCTGGCGAAGC 23429

RESULT 13
AX204969      3446 bp      DNA      linear      PAT 30-AUG-2001
LOCUS      AX204969/C
DEFINITION      Sequence 4 from Patent WO015180.
ACCESSION      AX204969
VERSION      AX204969.1 GI:15394251
KEYWORDS
SOURCE      Micromonospora carbonacea.
ORGANISM      Micromonospora carbonacea.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Micromonosporinae; Micromonosporaceae;
Micromonospora.
REFERENCE      1 (bases 1 to 3446)
AUTHORS      Stafla,A., Zazopoulos,E., Mercure,S. and Nowacki,P.
TITLE      Genetic locus for evernimicin biosynthesis
JOURNAL      Patent: WO 015180-A 4 02-AUG-2001;
Ecopia Biosciences Inc. (CA) ; Farnet, Chris (CA)
FEATURES
source      1..3446
              /organism="Micromonospora carbonacea"
              /db_xref="taxon:47853"
misc_feature      3..1037
                  /note="ORF 2 (positive strandedness) incomplete;
                  C-terminus only (N-terminus undetermined)."
```



gene  
CDS  
/translation="MAAGCAAAALFPWDRLTGAGSOAVASSLATAEAPLVPRAHAM  
LPKRLKPSFTSATSDLYEIRMOEQAVELVKELMSKVRPYDGPPIIRATQGEVVV  
ROINLQVNTAVNHLGHAIVLSNEDGLPMDITVPGEERTYRPNNOAPAASLYMDHANH  
LEAVNPGFGLNGILYLDHNERTLPJPGPYDVLVIRDAVEADGTLTLTRPSCDH  
MLVNNKERRPYOVAARKYRFRPVYACANRYKALRFADGIEPTQIGTGCGFLTPYOS  
ELMLMGERADIVDNFSRYKVGDSVLENPQASQIERPEWRDPIVTAPADYSYRFOR  
LTTIPQPTLPYVERDEFETIRTPAMTINQSYDPMKVIDITAKLTGETVTVKRNVEAPAA  
KPKDFPHLMHSHTHLTYFRVLERNGKAFTRDGLDVTYLIGPDVTYKIAMTMGPYT  
GQYLTGHQHLCHSSGCGMRIDIVA"  
/complement(1782. .2513)  
/gene="aur2H"  
complement(1782. .2513)  
/gene="aur2H"  
complement(1782. .2513)  
/gene="aur2H"  
complement(1782. .2513)  
/note="Aur2H"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAK61714.1"  
/db\_xref="GI:14486273"  
/translation="MSVVARELVAAGYRAGGTSGTGPGVGAGREPDGPGGAPCA  
VPGVAGVAGAGERCALRLLLSQPIGLLFAQAFVAMVYVYRTAPAPPOGDHAM  
WHSAPVPEPGRPLFLATLVAFYLLTPLLALGOVLRRCGRILVAGAGRSIPALD  
GPAATATGAVVTVPLVILYARFSPDNGELGAVYADAVTARLCAALALLAPFA  
CAPMAAAGASLRESGSLHSSS"  
complement(2510. .3367)  
/gene="aur2G"  
complement(2510. .3367)  
/gene="aur2G"  
complement(2510. .3367)  
/note="Aur2G"  
/codon\_start=1  
/transl\_table=11  
/product="oxygenase-like protein"  
/protein\_id="AAK61715.1"  
/db\_xref="GI:14486274"  
/translation="MSDTERVYLITGASRGLRAIAEVLNLEADTVVATARSIPALD  
GLVEFGRVYPVALDVTDRAYLAANAVEAEERIGRIDVLLNNAIGTLAAVEEVNT  
QVRDFDVFNEFGALVNTQAVLPYMKRQDSGHLQMSIAGVTYYPNLGLCASMALE  
AVSETLAEVAGFVKTALVEAGERRDMSGSMVATPKFAYDEVILARRHGISGAY  
AHILGDDPRKKGAEALTLTDEBPPLRVILGCGAADLAPVILRRLAGWEMDALART  
TDFPASEGP"  
complement(3364. .4314)  
/gene="aur2F"  
complement(3364. .4314)  
/gene="aur2F"  
complement(3364. .4314)  
/note="Aur2F"  
/codon\_start=1  
/transl\_table=11  
/product="cyclase/dehydroatase-like protein"  
/protein\_id="AAK61716.1"  
/db\_xref="GI:14486275"  
/translation="MTPGRLREVHEITVSAPAAVRYLIAEYVNMRIPEPTIYVQO  
VPEGGERIRITMANGEAENKMTNSRPLDAELRTPEQVTAAPVAAAGMITEE  
LSAQSRRLLHDIRAUNDODDSLAWIDEANDRSRSELAKKNVAAHAABELTS  
FEDVQIGSKADVIDFVNEADRWHERHPHATVRLBEDTPGLOELMDTRADGVSHT  
TTSRVYFAHRIARAQKQTLPALMTLHTGTWTEEBENGVAASSQHTVYINTDNIAR  
ILGESATYEDAKRYVHTALSTATLHAKAYAEQRAGR"  
complement(4349. .5134)  
/gene="aur2E"  
complement(4349. .5134)  
/gene="aur2E"  
complement(4349. .5134)  
/note="Aur2E"  
/codon\_start=1  
/transl\_table=11  
/product="ketoreductase-like protein"  
/protein\_id="AAK61717.1"  
/db\_xref="GI:14486276"  
/translation="MSDQGRVAAVVTGATSGIGLAVARLLAVOGRHVFIGARSAENVA  
ETVKLOEGEVLGAVYDVSTGVRAIMQADRFSGSVYVNNNGRSGGCVTADI  
AEEIMHDVITDINISVFPVREALITGSLKCDGRILYINISTAGRGVILGAPYAS  
KHGVGFKALGNEILAPRGITVNAVCPEYETTPMAQVIRGGTAAANDATDEALIEKQ  
AKVILGRISTEEVAGLVGLASDAASITSQALNVCGLGNF"  
/gene="aur2A"  
complement(6657. .7657)  
/gene="aur2A"  
complement(6657. .7657)

```

CDS
5386..6657
/gene="aur2A"
/note="Aur2A"
/codon_start=1
/transl_table=11
/product="B-ketoacyl-ACP synthase-like protein"
/protein_id="AAK61718.1"
/db_xref="GI:14486277"
/translation="MSRRVITIGVAVPGVGTKEFWSLTAGKTATRLSPDASP
FRSLIADADNPDLHGFTVAEMRLRPAOFALVSAEVRDSDIDPAELDPRITGV
TUGSANGCTTGLDYEYNVSRGSGTWOYDHTLAPHLFDYVPSMAEVAWRNAOC
PSMWSTGCTSGLDISGHAEVLIRESADVWITGTEAPISPTVACDAIKATPN
DDPGTASRPEDATRNQEVLGESAVLLEDESARKRAHVAELISGFASCMAYHMT
GLRADGVEMAEATVALDEARLDPSADYINAHSGTKQNDRHETAAKRSLGORYE
VPVSIKSMIGHSLGATGSSKVAASALAEYNTVPTANLHTPDLCDLDYPTLARE
ORDAVLSVSGFGFGFOSAMILTRPLGDAA"
6654..7916
/gene="aur2B"
6654..7916
/gene="aur2B"
/note="Aur2B"
/codon_start=1
/transl_table=11
/product="chain length factor-like protein"
/protein_id="AAK61719.1"
/db_xref="GI:14486278"
/translation="MTETTAARPGCTPLSPAAPVVTGIVTAENGLGTEAMMAVLR
GEGIRPYGRFATQYPAKLAGEVGEFPAEHPVSRLLPOTDHTRLALTADEALAD
AGVPAELPDSAGVITRASSAGFEFGOKELALMSGGGVSAVSAPMFYAVNTGO
ISIKNGLRGPGCVLYTEBOGGLDVAOARORLRGLRYVTGGVDSICPGMAHILA
GGELSTEDPARARPTPSAEDGHVAGGGALLVEDAAAREKGAELTGTAVGAST
FDPAPGSGGAPRAIAAEALALDAGLTPADVDVVFADAAGRDRADRAARALTLFGE
YGVVTAPEKMTGRLAAGASLIDLAALSLSDYIPIPTVTGPAEDCPVDLVSEVR
RPERLRALVYLAERGCGFNAMVRAAH"
7997..8500
/gene="aur2C"
7997..8500
/gene="aur2C"
/note="Aur2C"
/coiled_color=""
/codon_start=1
/transl_table=11
/product="oxigenase-like protein"
/protein_id="AAK61720.1"
/db_xref="GI:14486279"
/translation="MPIDPADFARTMAHVPYTIYTTVDAGORMGFTASFSSTLS
LDPLVLVCPAKTASCHDAEYTGDRFLVNVLAADIASFASGSHDKFTGSEMEPC
EFGPLGPDATARIACHTLHVLDGSDHSILVGRVESVSAGAEPLVYNNRLFTPGAS
PALASTS"
8524..9441
/gene="aur2D"
8524..9441
/gene="aur2D"
/note="Aur2D"
/codon_start=1
/transl_table=11
/product="cyclase-like protein"
/protein_id="AAK61721.1"
/db_xref="GI:14486280"
/translation="MTERSPSOALOEVAEDVAVYIOPDGMCLNAGLITRAGGRPAL
IDTATATARAALDAVARVPHAPGFYINHPGDDHFGNRLADEVITIAHATRA
EMDLAGLHTGLMPDVCCKDYGVLEPITLREPTITLHGLDIRAEVMSLSTATHRDSY
VWLPERRVLEAGDIVMSGATPFCLTGISISGLAVIEKRALGPRTIVSGHPVGSTEL
LDSTEAYLRVODLAGAGLAPLEAARADLGPFAALDSERLVNLHRAVAAEA
GAERGHLIDISELSEWDFHGGPLSCNA"

BASE COUNT      1370 a      3694 c      3132 g      1335 t
ORIGIN
Query Match      3.9%  Score 81.4; DB 1; Length 9531;
Best Local Similarity 54.1%; Pred. No. 4.8e-09;
Matches 216; Conservative 0; Mismatches 171; Indels 12; Gaps 2;

```

```

Oy 429 tacgaatgcatgtcccccaggaacctaccatcatcgtctccgtgtgacatgagatgtgtc 488
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1526 TACGACGGCACCCTCCCGCGCCGACCATCCGGGACACCAACCGAGGTGCTGTCGTC 1467
Oy 489 cgcttcgtgaacagcgagagagaacacccctcccaacagcgtccacttgcagcgtcttc 548
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1466 CG-----GCAGATCAGACAGCTCCAGGTGAACACCGCGTCCACTGTGACGCGGCGC 1413
Oy 549 tctgagctccctcttgatgtgtgtggtctgagacactaccacccgctggcgagatacaagat 608
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1412 GTGCTCTCCGACACAGACAGCGCTCCGATGACACCATCTGTCGGCGGCGCAACGTACG 1353
Oy 609 tactactaccaccaagaagcagcgtgccgacatgcttggtaacctatgacatgcatgtctc 668
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1352 TACCCCTACCCGACCAACACAGCGCGCGCGTGTGTGTGATCCACGACACACGCGCAC 1293
Oy 669 atcacgcgcgagaagcctacatggtgtgaagctgtgtgtctaatgatctcaagaccgcgt 728
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1292 CTGAGAGCCCGAGACAGCTCTCATGTGGGCTCGACAGGCTGTACTCTCTCACACCAAC 1233
Oy 729 gagatgccccctgaacctcccccagcgctacagcgagattgatatcccttgtctgact 788
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1232 GAGCGCACCGCTCCGCTGCGGACGCGTCCC-----TAGAGCTGCCGCTGCTCATCCGG 1179
Oy 789 gccaaagcatatacagcagcagcagcctctctctccacc 827
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1178 GACGCGCGGCTCGAGCGGACGCGACCCCTGCTCTACACC 1140

```

RESULT 15  
 LOCUS AP003233 137879 bp DNA linear PLN 25-MAY-2001  
 DEFINITION Oryza sativa genomic DNA, chromosome 1, PAC clone:P0037C04.  
 ACCESSION AP003233  
 VERSION AP003233.3 GI:14090356  
 KEYWORDS  
 SOURCE  
 ORGANISM

Oryza sativa (cultivar: Nipponbare) DNA, clone: P0037C04.  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (sites)  
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0037C04  
 JOURNAL Published Only in Database (2001) in press  
 REFERENCE 2 (bases 1 to 137879)  
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 TITLE Direct Submission

COMMENT  
 Submitted (19-FEB-2001) Takuji Sasaki, National Institute of  
 Agrobiological Resources, Rice Genome Research Program; Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@abrr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,  
 Tel:81-298-38-7441, Fax:81-298-38-7468)  
 On May 15, 2001 this sequence version replaced gi:1365490.

Genes were predicted from the integrated results of the following:  
 GENSCAN1.0, BLASTX2.0, BLASTX2.0 as well as SplicePredictor  
 (October 1998 version). The genomic sequence was searched against  
 NCBI Nonredundant protein database, nr  
 (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at  
 RGP. Protein homologues of the coding regions were searched against  
 NCBI Nonredundant protein database using BLASTN 2.0 with the  
 corresponding DBJ accession no. and RGP clone ID.  
 A gene with identity or significant homology to a protein is  
 classified based on the protein name to indicate the homology level  
 such as same name, 'putative-' and '-like protein'. A gene without  
 significant homology to any protein but with RST homology (covering  
 almost the entire length of partial sequence) is classified as an  
 'unknown' protein. A gene predicted with a gene prediction program  
 is classified as a 'hypothetical' protein.  
 The orientation of the sequence is from SP6 to T7 of the PAC clone.  
 This sequence of P0037C04 clone has an overlap with P0445D12 (DBJ:

AP003046) clone at the position 1 to 45,349 of 5' end. The sequence of this clone starts at the position 95,604 of P0445D12. This sequence of P0037C04 clone has an overlap with P0044F08 (DDBJ: AP002909) clone at the position 48,431 to 137,879 of 3' end. The sequence of this clone ends at the position 89,449 of P0044F08. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rtp.dna.affrc.go.jp/genomeseg.html>.

## FEATURES

```
source
    1. 137879
        /organism="Oryza sativa"
        /cultivar="Nipponbare"
        /db_xref="taxon:4530"
        /chromosome="1"
        /clone="P0037C04"
        join(905..1139,1595..1860)
        /gene="P0037C04.1"
        join(905..1139,1595..1860)
        /note="hypothetical protein"
        /codon_start=1
        /protein_id="BAB55515.1"
        /db_xref="GI:14090357"
        /translation="MAIWAARHGTAAQAPCGISLGMHGPDPADGCAVPAAGARTRRP
        MHGPIVGPGRAGPRAWAIVLPHKSLVLPQLSSSSKCYRHPRLULSKVYDERRR
        ROIHVCCGRRLPLSSORSSQPAVAVIMPPSPMRIRIGLSPTPPQQNDALHAA
        EVASA"
        complement(4956..8529)
        /gene="P0037C04.2"
        complement(4956..8529)
        /note="probably inactive due to frameshift(s) in CDS
        pseudogene, similar to gag-pol polyprotein"
        /pseudo
        join(10530..10652,12160..12528)
        /gene="P0037C04.3"
        join(10530..10652,12160..12528)
        /gene="P0037C04.3"
        /note="hypothetical protein"
        /codon_start=1
        /protein_id="BAB55516.1"
        /db_xref="GI:14090358"
        /translation="MODIPHEERLLDYREKSRDQDKLPWGVQSMATDKTKPN
        LISTIYVLVHREDSRVLLNGDVPKGRSLVYCAAGAEALTYMTAVSGAP
        GALSLASGGSVPRVRRWRYPTGGFLFPDAPWRASGGSVTVVHKLPPELEDT
        TAA"
        complement(15496..16914)
        /gene="P0037C04.4"
        complement(15496..16914)
        /gene="P0037C04.4"
        /note="contains EST C28475(C61239)
        similar to Oryza sativa chromosome 1, P0698A04.16
        unknown protein"
        /codon_start=1
        /protein_id="BAB55517.1"
        /db_xref="GI:14090359"
        /translation="MADRRPPCVLLERVVRVEAAGLTSGCASRDADVAATIEAGW
        SRSRVKMGSAEEMERRAPSVKPVAFADPQASSLNLHLPAPARTLLIGETISGT
        HKGIYVIAHRCYLLYDASNNHTA1IPVPDVFVPLGRSAVLSAAGDDDDYLIAD
        IVTSCSRGIPNLPALKATIFAVKNGEIOSSIPHLPLPILGPTFFHIDTAFSF
        AGTIFWVLDLKGILDEILSDQGRVLFVLRHCIDAHDRHCFSEF2HRSIGRV
        SGATKFLALIGCEASCPANVRLKWTSLSPDFKMKKEETLITVGDIMKSSPOMG
        LPHVLPSPILGVNEDGIMYVNLNVKEEPIRILNEPDSLGIOYPRKANTIRDM
        ONKVLSSKTKISKPTFRNFTMTFTLASDFSAYIQDQVNLNLTATLATHTFSLADHP
        NKLVCFTLFTQTLILFVTS5"
        complement(join(18962..19079,19226..19636,19997..20204,
        20474..20575,20871..20948,21217..21321,21330..21552))
        /gene="P0037C04.5"
        complement(join(18962..19079,19226..19636,19997..20204,
        20474..20575,20871..20948,21217..21321,21330..21552))
        /gene="P0037C04.5"
        /note="hypothetical protein"

gene
    /codon_start=1
    /protein_id="BAB55518.1"
    /db_xref="GI:14090360"
    /translation="WVAALVGEQVDVADPTNGMGRRRPPYLRRGGGCGGRRPPD
    VAGGVVEKASRSRGYPDPMPGPDPAVGGGCGYRRIRLRPLSGGGCGGGRRA
    ERTQSGDEKREKREKREDREEREKREKRLVDEYISDEALWSTITNRSOMGRGV
    VLLKTRTSODRDSVAOKODPPASDSDRDVDGAGRAPSKIPARTETIWPPTPHAS
    DARGNARYROMIEPVGAHLGRLLPVAARLLPPAASRRPPPPPPADALARIY
    VVAASSSSSRLPSSSPSPAELSLAGRARLISRRPPRSPSSPLSSPAP
    AAELSFSLAGHARRRRRPPSPSSPAPAVLSAASLASSNQTLSKKRRCPILRY
    PTNQTHERVRRMSRYAAAEGR"
    join(25223..25413,25596..26220)
    /gene="P0037C04.6"
    join(25223..25413,25596..26220)
    /gene="P0037C04.6"
    /note="hypothetical protein"
    /codon_start=1
    /protein_id="BAB55519.1"
    /db_xref="GI:14090361"
    /translation="MRLPRLGMKKNGKRTTMDHGTAKKTRSVVAKKTEAGLCIDQ
    ODVEERTKMTYSIDSDLEACPSCNEPIGDIRCPLEKTVLAAMAPCKFRASGME
    VGFETRLSHEASCTHAPCGCPDCTYIGLILYNNILDEHATDAVYAMSLGTVTV
    HSKPENVLLHKGTRGKRVFLNLNGDVLISGRSLSVSGVPPPTANCLLYKTELA
    ADGPCTGVILSLASGTVPCVRLGEFNKAFLEVPDPSYSGSSDFTVSVTVLI"
    complement(join(27855..28310,29207..29638,29878..29901))
    /gene="P0037C04.7"
    complement(join(27855..28310,29207..29638,29878..29901))
    /gene="P0037C04.7"
    /note="hypothetical protein"
    /codon_start=1
    /protein_id="BAB55520.1"
    /db_xref="GI:14090362"
    /translation="MMWLRLROANRRGAAAEQAVASGEPPDGRARRPCGPPYAR
    SWRRPVSCKLASAERQCAVYAASSPAMCOTBGRGGGCPALQILRRHAAKVASR
    LARCBEGARRRQQAAGHSGSGSNSTGTSOPTVYEDLKEIGLARARRRA
    AAEQAVASGEPPDGRARRPCGPPYARHARRMPVSCKLASERGAOVAASPEPMQ
    TEGRGGGCPALQILRRHAAKVASRSIARREGARRRGAQAQRRGSGSNST
    TOSOPTVYEDLKEIGLKVTSNTNR"
    complement(join(30803..31498,31608..31763))
    /gene="P0037C04.8"
    complement(join(30803..31498,31608..31763))
    /gene="P0037C04.8"
    /note="hypothetical protein"
    /codon_start=1
    /protein_id="BAB55521.1"
    /db_xref="GI:14090363"
    /translation="MDNTGTPKRTKMSENKAAQODKDETTKLTYSIERDAECPICF
    PFEDRVYMKNGHAAAGSCYAKAMTMCPDCEIPICINICRPLEKYLAAASAPCRSTS
    AYMRILRASGCTEIVATERRHNEASCPHAPCPEDGNNVGHLLYSIHIDEHATDA
    AVATGCLRGCTGVTVLHKRSPFHVLLHRRGSRVLLNGDVLISGRSLSYCTSP
    PLPNCBELKXIGLAVSRAPGELGLSMGSTVYCVRLBGFDAKATLFPVDSYWGSSGT
    ISVTVHL"
    join(32915..33391,33775..34458)
    /gene="P0037C04.9"
    join(32915..33391,33775..34458)
    /gene="P0037C04.9"
    /note="hypothetical protein"
    /codon_start=1
    /protein_id="BAB55522.1"
    /db_xref="GI:14090364"
    /translation="MSGGVGRDRIAPAGTARRRRRPPPTTGSPESVYLEGWSLSP
    AYPARRRRHDEGYSRRGRSRSSPPARRRRRSPRSRSRTSYCEDSCCHTR
    BSPRDDEDDDDDDGGSYHRRPADNETFVRLDYGADDGIFRCGCTAMLSPPITEC
    ANGDVIERCSTDDGGAVRCRCGTWELARSAIHLKLCIRFACKNNRYGSPSLPR
    OMDHDEHLSCHDEPCFPIRRGCFAGAADSLRLHRTAAGWGRGLVAAAGEAAYVVO
    PTLIADGRIPIHLSTCRERGGGCTAASMC1IRPDHYAGAEETVYERPAACRLDM
    QAAVGTSLRYGKMDAVOARVYVDDMLLRQDAAQVYVRRATSAAGANNNN"
    complement(join(35545..35965,38859..39226))
    /gene="P0037C04.10"
    complement(join(35545..35965,38859..39226))
    /gene="P0037C04.10"
    /product="putative Bowman Birk trypsin inhibitor"
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 20:28:34 ; Search time 448.02 Seconds

(without alignments)  
8028.516 Million cell updates/sec

Title: US-09-656-640A-3

Perfect score: 2095

Sequence: 1 cagctcgctactactctc.....gaataagttggtgcttaa 2095

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq\_032802.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2095	100.0	2095	22	AAF82586
2	2095	100.0	2095	22	AAF30029
3	1946.2	92.8	1958	22	AAF30028
4	1944.6	92.8	1958	22	AAF82585
5	528.6	25.2	1959	14	AA047790
6	517.4	24.7	2110	21	AA047790
7	460.4	22.0	1791	20	AA227601
8	460.4	22.0	1791	21	AA50019
9	460.4	22.0	1791	21	AA51314

10	458.8	21.9	1791	20	AA225727	Stachybotrys chart
11	300.4	14.3	2067	20	AA225735	Stachybotrys chart
12	300.4	14.3	2067	20	AA227609	Stachybotrys pheno
13	300.4	14.3	3676	20	AA225728	Stachybotrys chart
14	300.4	14.3	3677	20	AA227602	Stachybotrys pheno
15	300.4	14.3	3677	21	AA50018	Stachybotrys chart
16	300.4	14.3	3677	21	AA51313	Stachybotrys chart
17	288	13.7	2063	21	AA50021	Curtularia pallasc
18	288	13.7	2063	21	AA51316	Curtularia pallasc
19	283	13.5	2905	21	AA50020	Bipolaris spicifer
20	283	13.5	2905	21	AA51315	Bipolaris spicifer
21	95.8	4.6	109519	22	AA508693	Micromonospora DNA
22	87.4	4.2	858	21	AA50022	Amerosporium atrum
23	87.4	4.2	858	21	AA51317	Amerosporium atrum
24	49.6	2.4	1533	22	AA66018	C glutamicum codin
25	49.6	2.4	1533	22	AA66018	Cornebacterium gl
26	49.6	2.4	349980	22	AA668527	C glutamicum codin
27	47.2	2.3	523	21	AA43363	Rat secreted expe
28	47.2	2.3	523	21	AA43363	Rat secreted expe
29	44.8	2.1	10732	21	AA10594	Gene encoding a su
30	44.8	2.1	61313	23	AA559545	Propionibacterium
31	43.6	2.1	68356	22	AA67283	Human immune/haema
32	43.6	2.1	68356	22	AA67283	Human immune/haema
33	41.2	2.0	2012	22	AA57385	Human skeletal mus
34	40.8	1.9	600	22	ABA59948	Human foetal liver
35	40.8	1.9	600	22	ABA28371	Probe #6837 for ge
36	40.8	1.9	600	22	AAK08219	Human brain expres
37	40.8	1.9	600	22	AAK34097	Human bone marrow
38	40.8	1.9	600	22	AA139820	Probe #8506 used t
39	40.4	1.9	15772	22	AAK83220	Human immune/haema
40	39.4	1.9	2071	15	AA058121	pH 2.5 acid phosph
41	39.4	1.9	2071	15	AA058121	A. niger pH 2.5 ac
42	38.8	1.9	523	21	AA081723	P. meningitidis pa
43	38.8	1.9	20757	20	AAK20599	Polynucleotide seq
44	38.8	1.9	23457	22	AAK74866	Human immune/haema
45	38.8	1.9	31405	22	AAK67293	Human immune/haema

## ALIGNMENTS

RESULT 1	AAF82586	standard; DNA; 2095 BP.
ID	AAF82586	
AC	AAF82586	
XX		
DT	18-JUN-2001	(first entry)
DE	Stachybotrys chartarum phenol oxidase B gene.	
XX		
KW	Stachybotrys phenol oxidase B; spob; phenol oxidising enzyme;	
KW	detergent; paper production; pulp production; textile; food industry;	
KW	bleaching; ds.	
OS	Stachybotrys chartarum.	
XX		
XX		
FT	key	Location/Qualifiers
FT	CDS	77..2016
FT	exon	/tag= a
FT	exon	/product= "Stachybotrys phenol oxidase B"
FT	intron	77..267
FT	intron	/number= 1
FT	exon	268..319
FT	exon	/number= 1
FT	intron	320..869
FT	intron	/number= 2
FT	intron	870..952
FT	exon	/number= 2
FT	exon	953..1692
FT	exon	/number= 3
FT	intron	1693..1745
FT	intron	/number= 3

FT exon 1746..2013  
PT /number= 4  
PN WO200121748-A1.  
XX 29-MAR-2001.  
PD 06-SEP-2000; 2000WO-EP08840.  
PE 22-SEP-1999; 99EP-0203120.  
XX  
XX (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
PA (HIND-) HINDUSTAN LEYER LTD.  
PI Convents D, Doornink M, De Vries CH, Wang H;  
XX  
XX WPI: 2001-273462/28.  
DR P-PSDB: AAB81505.  
XX  
PT New detergent compositions comprising a phenol oxidizing enzyme useful  
PT in detergent or cleaning compositions, fiber treatment, processing,  
PT finishing or production, paper and pulp production, or in starch  
PT processing applications  
PS  
PS Example 3; Fig 3; 46pp; English.  
XX  
XX The present sequence encodes a Stachybotrys chartarum phenol oxidizing  
CC enzyme. The invention relates to detergent compositions comprising one or  
CC more surfactants and a phenol oxidizing enzyme having at least 688  
CC identity to the Stachybotrys chartarum phenol oxidizing enzyme. Phenol  
CC oxidizing enzymes may be used in the detergent, paper, pulp, textile and  
CC food industries. They are used for preventing the transfer of dyes in  
CC solution from one textile to another during detergent washing, or in  
CC modifying the colour associated with dyes and coloured compounds having  
CC different chemical structures, such as in pulp and paper bleaching,  
CC bleaching the colour of stains on fabric and in detergent and textile  
CC applications.  
XX  
XX Sequence 2095 BP; 437 A; 618 C; 510 G; 530 T; 0 other;

Query Match	Similarity	100.0%	Score 2095;	DB 22;	Length 2095;
Best Local	Similarity	100.0%	Prod. No. 0;		
Matches 2095;	Conservative	0;	Mismatches	0;	Indels
					Gaps
					0;
QY	1	cagctcgagctactactcgcgtctctctttggaacaatcaactacacatcgcttcctca	60		
Db	1	cagctcgagctactactcgcgtctctctttggaacaatcaactacacatcgcttcctca	60		
QY	61	attcaaacgatacaacatgatacagccaagctatcgagagcgctgctctgagcttgctg	120		
Db	61	attcaaacgatacaacatgatacagccaagctatcgagagcgctgctctgagcttgctg	120		
QY	121	gctcgcgagcgagctctgctgcgagtcgagacgcgattgctgtgtgcatacgacaagatcgctc	180		
Db	121	gctcgcgagcgagctctgctgcgagtcgagacgcgattgctgtgtgcatacgacaagatcgctc	180		
QY	181	cggtctccaccaaagagcgacagcgagctgagtcctccctcccgagctctgaagaagtgctct	240		
Db	181	cggtctccaccaaagagcgacagcgagctgagtcctccctcccgagctctgaagaagtgctct	240		
QY	241	gccgatacctctctcttgaaagcgagcgcccaagtaagtaagtaacttctatagcttaagacagcca	300		
Db	241	gccgatacctctctcttgaaagcgagcgcccaagtaagtaagtaacttctatagcttaagacagcca	300		
QY	301	acgttgctaatcatgcaatagcaatagcgtaccgctcccaaccccaacaccttgagagagacatctgta	360		
Db	301	acgttgctaatcatgcaatagcaatagcgtaccgctcccaaccccaacaccttgagagagacatctgta	360		
QY	361	acgagatgtagagatgaagccctctctccaccagatacctctgatactgagacgagccaaca	420		
Db	361	acgagatgtagagatgaagccctctctccaccagatacctctgatactgagacgagccaaca	420		

QY	421	tgrtttgatagcagatgacatgtccccaagacctaacaatcgtttccctcgtgacatgaga	480
Db	421	tgrtttgatagcagatgacatgtccccaagacctaacaatcgtttccctcgtgacatgaga	480
QY	481	gtgtgttcgcgttcgtgaaacagcggagagaacacctctccaaacagcgttcaactgtgacg	540
Db	481	gtgtgttcgcgttcgtgaaacagcggagagaacacctctccaaacagcgttcaactgtgacg	540
QY	541	gctctttctcgcagctcccttctgattggttggtgcgtgaggaacataccaagccctgacgatt	600
Db	541	gctctttctcgcagctcccttctgattggttggtgcgtgaggaacataccaagccctgacgatt	600
QY	601	acaagattactactataccccaacagaagcaggtctcccgacatgtctttgtgtacatgacatg	660
Db	601	acaagattactactataccccaacagaagcaggtctcccgacatgtctttgtgtacatgacatg	660
QY	661	ccatgtccatcaacacgcgcgagagacgcctacatggtctcaagctcgtgtctacatgataccag	720
Db	661	ccatgtccatcaacacgcgcgagagacgcctacatggtctcaagctcgtgtctacatgataccag	720
QY	721	accgcgctcgtgagatgccccctgaacctccccaagcgtcgaagcgagattgatatccctctgg	780
Db	721	accgcgctcgtgagatgccccctgaacctccccaagcgtcgaagcgagattgatatccctctgg	780
QY	781	tctctgacgtcccaacgcgataacacgcgacacgcacctctctcccaacaaatgagaggttt	840
Db	781	tctctgacgtcccaacgcgataacacgcgacacgcacctctctcccaacaaatgagaggttt	840
QY	841	ccaagctctcgtggtgtgacgttataccaagtgtgtaagttgagcccatgtgagatgctccaagtc	900
Db	841	ccaagctctcgtggtgtgacgttataccaagtgtgtaagttgagcccatgtgagatgctccaagtc	900
QY	901	ctaagagatcatcgatgtatctggaatatgtgtgcaatgacctaaacagtgacctataccaagaaagtgca	960
Db	901	ctaagagatcatcgatgtatctggaatatgtgtgcaatgacctaaacagtgacctataccaagaaagtgca	960
QY	961	gacctgagcctatgtctcaacgltgcagccgcgcgaagtacacgctctccgctctccaacgctgc	1020
Db	961	gacctgagcctatgtctcaacgltgcagccgcgcgaagtacacgctctccgctctccaacgctgc	1020
QY	1021	cgctccaaagcctcttcgcctctctgtatctctgtacacctctgaaggtctcaagaaacagaccttc	1080
Db	1021	cgctccaaagcctcttcgcctctctgtatctctgtacacctctgaaggtctcaagaaacagaccttc	1080
QY	1081	cttccaaaggtctatctgcgcgtctgacgctgagctcgtctcgaagggccctgtcttgacaatgtacaactct	1140
Db	1081	cttccaaaggtctatctgcgcgtctgacgctgagctcgtctcgaagggccctgtgttgacaatgtacaactct	1140
QY	1141	gtacatctctctatctgcccgaagcgcgtcggaaggtgtgtgtatcgaactcttccacctctgcgtcgca	1200
Db	1141	gtacatctctctatctgcccgaagcgcgtcggaaggtgtgtgtatcgaactcttccacctctgcgtcgca	1200
QY	1201	gtccatctgatatctgcgaacctctccgtgtgcgcggtctccggtgtcttgagccttgagtttga	1260
Db	1201	gtccatctgatatctgcgaacctctccgtgtgcgcggtctccggtgtcttgagccttgagtttga	1260
QY	1261	taaacatctgacaaggtctatctgcgattcgttcgtgtgaatgcctcttgatcgcgcgcgaacttc	1320
Db	1261	taaacatctgacaaggtctatctgcgattcgttcgtgtgaatgcctcttgatcgcgcgcgaacttc	1320
QY	1321	tgaagttgcctgcgaacctctcgagatgttcccttccccaagggccggaactgtgagaccgcgc	1380
Db	1321	tgaagttgcctgcgaacctctcgagatgttcccttccccaagggccggaactgtgagaccgcgc	1380
QY	1381	aaaccccaactatgacaagacattccacctcgcgcgtgtgataatgtgacaagtgagacaatcaa	1440
Db	1381	aaaccccaactatgacaagacattccacctcgcgcgtgtgataatgtgacaagtgagacaatcaa	1440
QY	1441	cggagttacactctctcggatgtctgcgaagaacgltcgtctcgcgaatgtgtcccgcgacactgtc	1500
Db	1441	cggagttacactctctcggatgtctgcgaagaacgltcgtctcgcgaatgtgtcccgcgacactgtc	1500

QY	1081	cttcgaagtcattgcgcgtctaacgctgagctgcagctcttgagggccctgtcttgacaactggaacattctc	1140
Db	1081	cttcgaagtcattgcgcgtctaacgctgagctgcagctcttgagggccctgtcttgacaactggaacattctc	1140
QY	1081	cttcgaagtcattgcgcgtctaacgctgagctgcagctcttgagggccctgtcttgacaactggaacattctc	1140
Db	1081	cttcgaagtcattgcgcgtctaacgctgagctgcagctcttgagggccctgtcttgacaactggaacattctc	1140
QY	1141	gtacatctctatctgacgagcgctgaggagtgctgttatcgactctccacacttcgcctgctgagca	1200
Db	1141	gtacatctctatctgacgagcgctgaggagtgctgttatcgactctccacacttcgcctgctgagca	1200
QY	1201	gtccatcgatatccgcgaacctctccctgcgtctgagccggtctccggtgtgttgagccttgagtttga	1260
Db	1201	gtccatcgatatccgcgaacctctccctgcgtctgagccggtctccggtgtgttgagccttgagtttga	1260
QY	1261	taaaccttgacaaagtcactgtgattctgtctgttgtaagctcttgagtgctgcgccgaacattc	1320
Db	1261	taaaccttgacaaagtcactgtgattctgtctgttgtaagctcttgagtgctgcgccgaacattc	1320
QY	1321	tgaagctgacctcaacctccgagatgtctcttcccccgaaggcggaactggaacccgcg	1380
Db	1321	tgaagctgacctcaacctccgagatgtctcttcccccgaaggcggaactggaacccgcg	1380
QY	1381	aaaccccactgatgacgagacttcaaccttcgcgcgtgtgtaattgacacagtgagacaatcaa	1440
Db	1381	aaaccccactgatgacgagacttcaaccttcgcgcgtgtgtaattgacacagtgagacaatcaa	1440
QY	1441	cggagctaacctctctcgatgtcgagaaacgctctgcctccgcgaattgtgcgccgcgacactgtc	1500
Db	1441	cggagctaacctctctcgatgtcgagaaacgctctgcctccgcgaattgtgcgccgcgacactgtc	1500



Db 421 tgggtggaatacgaatgcalgtcccaaggactacaatcagltccctcgtggaactgaga 480  
 Qy 481 gttgtccgccttcgtgaacagcgagaaacacactctcccaacagcgccacttgaca 540  
 Db 481 ggtgttcgcgcttcgtgaacagcgagaaacacactctcccaacagcgccacttgaca 540  
 Qy 541 gctcttctctcgaagctccctttgagltgttggtcgtgaggaacaactaacgacctgagat 600  
 Db 541 gctcttctctcgaagctccctttgagltgttggtcgtgaggaacaactaacgacctgagat 600  
 Qy 601 acaaggaattactactaccccccaacagcgagcgcccgacgtgtttgttaccatgaccatg 660  
 Db 601 acaaggaattactactaccccccaacagcgagcgcccgacgtgtttgttaccatgaccatg 660  
 Qy 661 ccatgtccacaccccgcgagaaacgacctacatgtgtcagcgtgtgtctacatgacatccag 720  
 Db 661 ccatgtccacaccccgcgagaaacgacctacatgtgtcagcgtgtgtctacatgacatccag 720  
 Qy 721 acccggtcgtgagatgacctgaaactcccccagcggtacgagatttgaatcccttgg 780  
 Db 721 acccggtcgtgagatgacctgaaactcccccagcggtacgagatttgaatcccttgg 780  
 Qy 781 tctctactgcgaacgatacaacgacgacgacactctcttcccaacaaatgagagattt 840  
 Db 781 tctctactgcgaacgatacaacgacgacgacactctcttcccaacaaatgagagattt 840  
 Qy 841 ccaggtctcgtgagatgacgttatctcaagtgtgagtgtgagccatgtgatatgtcttcagtc 900  
 Db 841 ccaggtctcgtgagatgacgttatctcaagtgtgagtgtgagccatgtgatatgtcttcagtc 900  
 Qy 901 ctagaagatcgaatgatagaatgtgtgcatgctcttaaccagtgctatacagaacggtca 960  
 Db 901 ctagaagatcgaatgatagaatgtgtgcatgctcttaaccagtgctatacagaacggtca 960  
 Qy 961 gctcttgccatcgtcgaacgctgacgacgacgacgacgacgacgacgacgacgacgacgac 1020  
 Db 961 gctcttgccatcgtcgaacgctgacgacgacgacgacgacgacgacgacgacgacgacgac 1020  
 Qy 1021 cgtctcaacgctcttcgctctgtatcttgtactacactcgtgagatcagagacgacactcc 1080  
 Db 1021 cgtctcaacgctcttcgctctgtatcttgtactacactcgtgagatcagagacgacactcc 1080  
 Qy 1081 ctctcaagatcattgcgcgtgacggtgtgtctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1140  
 Db 1081 ctctcaagatcattgcgcgtgacggtgtgtctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1140  
 Qy 1141 gtacatctctatggcgagcgctgagaggttgttatgagctctccacacttcgcgtgagcga 1200  
 Db 1141 gtacatctctatggcgagcgctgagaggttgttatgagctctccacacttcgcgtgagcga 1200  
 Qy 1201 gtccatcgtatccgcaacactcctcgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260  
 Db 1201 gtccatcgtatccgcaacactcctcgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260  
 Qy 1261 taacactgacaaggtcgaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1320  
 Db 1261 taacactgacaaggtcgaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1320  
 Qy 1321 tgaagtgcctgcgaacccctccgagatggttcccttcccgagggcggaacatggaaccccg 1380  
 Db 1321 tgaagtgcctgcgaacccctccgagatggttcccttcccgagggcggaacatggaaccccg 1380  
 Qy 1381 aaaccccaactgatacgaagacttcaacttcgcgcgtgtgtaatacgaacgtggaacatca 1440  
 Db 1381 aaaccccaactgatacgaagacttcaacttcgcgcgtgtgtaatacgaacgtggaacatca 1440  
 Qy 1441 cggagttacccttcgtgagatgtaagaacacgctgtcgcgaatgtgccccgcgacatgt 1500  
 Db 1441 cggagttacccttcgtgagatgtaagaacacgctgtcgcgaatgtgccccgcgacatgt 1500  
 Qy 1501 tgaagtgcgcgacgtggaacaaactcgaacggttggaatacactgttccatcaact 1560  
 Db 1501 tgaagtgcgcgacgtggaacaaactcgaacggttggaatacactgttccatcaact 1560

Qy 1561 cgttgaattccagatccttctcgttccacatgtcccggtgagatcgaagccttaatgagctgc 1620  
 Db 1561 cgttgaattccagatccttctcgttccacatgtcccggtgagatcgaagccttaatgagctgc 1620  
 Qy 1621 tgggtcgaagatgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1680  
 Db 1621 tgggtcgaagatgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1680  
 Qy 1681 cgtctctcccgtaagttcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1740  
 Db 1681 cgtctctcccgtaagttcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1740  
 Qy 1741 acaagtgtgtctacatgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1800  
 Db 1741 acaagtgtgtctacatgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1800  
 Qy 1801 gctgtcttcaatgtcaactgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1860  
 Db 1801 gctgtcttcaatgtcaactgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1860  
 Qy 1861 cccatgagacgtctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1920  
 Db 1861 cccatgagacgtctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1920  
 Qy 1921 ggtgtactcagcagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1980  
 Db 1921 ggtgtactcagcagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1980  
 Qy 1981 taagcccaagcgtgatacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2040  
 Db 1981 taagcccaagcgtgatacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2040  
 Qy 2041 tggagacagatatacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2095  
 Db 2041 tggagacagatatacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2095

RESULT 3  
 AAF30028  
 ID AAF30028 standard; DNA; 1958 BP.  
 XX  
 AC AAF30028;  
 XX  
 DT 23-APR-2001 (first entry)  
 XX  
 DE Stachybotrys chartarum phenol oxidase B spob gene.  
 XX  
 KW Phenol oxidizing enzyme; phenol oxidase B; spob gene; bleach;  
 KW pulp; paper; textile; detergent; ss.  
 OS  
 XX Stachybotrys chartarum.  
 XX  
 FH  
 FH Key  
 FT CDS  
 FT 13..1952  
 FT /\*tag= a  
 FT /note= "contains introns"  
 FT 1..203  
 FT /\*tag= b  
 FT 204..255  
 FT /\*tag= c  
 FT 256..805  
 FT /\*tag= d  
 FT 806..888  
 FT /\*tag= e  
 FT 889..1628  
 FT /\*tag= f  
 FT 1629..1681  
 FT /\*tag= g  
 FT 1682..1958  
 FT /\*tag= h  
 PN US6168936-B1.





CC	food industries. They are used for preventing the transfer of dyes in solution from one textile to another during detergent washing, or in modifying the colour associated with dyes and coloured compounds having different chemical structures, such as in pulp and paper bleaching, CC bleaching the colour of stains on fabric and in detergent and textile applications.
CC	
xx	
SQ	Sequence 1958 BP; 394 A; 592 C; 491 G; 481 T; 0 other:
Query Match	92.8%; Score 1944.6; DB 22; Length 1958;
Best Local Similarity	99.8%; Pred. NO. 0;
Matches 1947; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Dy	67 aacgaatcaaatgatacagcaaccaagatcatcgaaaccgttgctctgggccttgcgtgatccg 126 
Dd	3 atcatcaacaatgatacagcaaccaagatcatcgaaaccgttgctctgggccttgcgtgatccg 62 
OY	127 cggcacgctctgtcgatgccagatccgtttgcttgatcatgcatacagaacatgaccttcgatt 186 
Dd	63 cgycagactctgtcgatgccagatccgtttgcttgatcatgcatacagaacatgaccttcgatt 122 
OY	187 caaccaaggagcagcaagcagacttgagatctccctctggtccttgttaacgaagtgcctctgcgat 246 
Dd	123 cacccaaggagcagcaagcagacttgagatctccctctggtccttgttaacgaagtgcctctgcgat 182 
OY	247 ccctcctcttgaaagcgcaccaagatgataaagctaacattctlaagctctacgaagaacacacgttta 306 

5  
6

5

2  
4

28

5

2.

1

3

1

2.

2

1

1

•

•

2

```

QY 967 gctatgctcaacgtgcaacgcgcgaagtaaccgcttcctcctcaacgctgcgtc 1026
    |||
Db 903 gctatgctcaacgtgcaacgcgcgaagtaaccgcttcctcctcaacgctgcgtc 962
QY 1027 acgctcttcgcctgctatcctgtacacctcgtgagatcaagagaccagcttcctca 1086
    |||
Db 963 acgctcttcgcctgctatcctgtacacctcgtgagatcaagagaccagcttcctca 1022
QY 1087 ggtcattgcgcgcgaaggtgctgtgttgaggccctgtttgaaactgaaactctgtacat 1146
    |||
Db 1023 ggtcattgcgcgcgaaggtgctgtgttgaggccctgtttgaaactgaaactctgtacat 1082
QY 1147 cctcatgscgcgcgcgtgagaggtgtgtatcgacctcccaaccttcgcgcgcgcagctccat 1206
    |||
Db 1083 cctcatgscgcgcgcgtgagaggtgtgtatcgacctcccaaccttcgcgcgcgcagctccat 1142
QY 1207 cgatatacgcaacctctcgtgtcgtgacggtctcgggtgttgagcctgagttgtataaac 1266
    |||
Db 1143 cgatatacgcaacctctcgtgtcgtgacggtctcgggtgttgagcctgagttgtataaac 1202
QY 1267 tgacaaggtcagtgagttgtgtgttgatgaagtccttgagtgccgcgcgcgcagctctgaggt 1326
    |||
Db 1203 tgacaaggtcagtgagttgtgtgttgatgaagtccttgagtgccgcgcgcgcagctctgaggt 1262
QY 1327 gctcgcgaacctcgcgaagatgttcccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1386
    |||
Db 1263 gctcgcgaacctcgcgcgaagatgttcccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1322
QY 1387 caactgatacgaggaacttcaaccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1446
    |||
Db 1323 caactgatacgaggaacttcaaccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1382
QY 1447 taacctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1506
    |||
Db 1383 taacctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1442
QY 1507 ctgagcgaacttgagaaacaactcccaacgcgttgagactcaaccttcgcgcgcgcgcgcgcgcgcgcgc 1566
    |||
Db 1443 ctgagcgaacttgagaaacaactcccaacgcgttgagactcaaccttcgcgcgcgcgcgcgcgcgcgcgc 1502
QY 1567 cttcgcgaactccttcctcgttcgaactgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1626
    |||
Db 1503 cttcgcgaactccttcctcgttcgaactgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1562
QY 1627 caaggaatgttgcgtgctgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1686
    |||
Db 1563 caaggaatgttgcgtgctgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1622
QY 1687 ttcccgtaagtctcgcgttccttaactaactggtttcaactaactgcttaactaactaactaactaact 1746
    |||
Db 1623 ttcccgtaagtctcgcgttccttaactaactggtttcaactaactgcttaactaactaactaactaact 1682
QY 1747 ggtgtctacatggttgcacgtgcacaacactgataccagagaccagacatgtagtgcgt 1806
    |||
Db 1683 ggtgtctacatggttgcacgtgcacaacactgataccagagaccagacatgtagtgcgt 1742
QY 1807 ttcaatgtacatggttgcgtgactatggtctaaactaactaactaactaactaactaactaactaact 1866
    |||
Db 1743 ttcaatgtacatggttgcgtgactatggtctaaactaactaactaactaactaactaactaactaact 1802
QY 1867 gggcctctctgagagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1926
    |||
Db 1803 gggcctctctgagagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1862
QY 1927 ttcaagcgaacttgccactcaactgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1986
    |||
Db 1863 ttcaagcgaacttgccactcaactgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1922
QY 1987 caggcgtgatagtatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2046
    |||
Db 1923 caggcgtgatagtatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1982

```

```

RESULT 5
AAQ47790
ID AAQ47790 standard; DNA; 1959 BP.
XX
AC AAQ47790;
XX
DT 24-FEB-1994 (first entry)
XX
DE Bilirubin oxidase gene.
XX
KW Bilirubin; oxidase; analytical; BO; expression vector; PCR;
KW polymerase chain reaction; ss.
OS Myrothecium verrucaria.
XX
FH Key Location/Qualifiers
FT CDS 66..1784
FT /*tag= a
FT /*product= BO
FT sig_peptide 66..179
FT /*tag= b
FT mat_peptide 180..1781
FT /*tag= c
XX
PN JP05199882-A.
XX
PD 10-AUG-1993.
XX
PE 24-JAN-1992; 92JP-0034126.
XX
PR 24-JAN-1992; 92JP-0034126.
XX
PA (AMANO ) AMANO PHARM KK.
XX
DR WPI; 1993-284681/36.
DR P-PSDB; AAR40843.
XX
PT Bilirubin oxidase prepn. useful as an analytical enzyme - by
PT culturing Bilirubin oxidase in transformant culture
XX
PS Claim 1; Page 29-32; 32pp; Japanese.
XX
CC The sequence encodes bilirubin oxidase. The protein produced has a
CC 38 amino acid signal peptide which is removed to give the mature
CC protein (AAR40843).
XX
SQ Sequence 1959 BP; 457 A; 549 C; 436 G; 517 T; 0 other;
XX

Query Match 25.2%; Score 528.6; DB 14; Length 1959;
Best Local Similarity 62.0%; Pred. No. 6; 7e-139;
Matches 1058; Conservative 0; Mismatches 474; Indels 175; Gaps 7;
QY 303 gttgctaactatgacagtcacgctcccaaccccaacacactgagagagacatctgtactac 362
    |||
Db 237 gttgaagcagcccgcttgactgttaaccaatcctgtgaatggaagaagatctgtactat 296
QY 363 gagaatgagatgagccctcctccacacagatcacctgatactgtgagccgcgaacatg 422
    |||
Db 297 gaggctgagatacagcccttcaactcaactcaggttaacctgactgtgacatcgtgactgc 356
QY 423 gttgatacagatgagccttccccaaggaactacatcatcgttccctcgtgacatgagat 482
    |||
Db 357 gtcggatcagatgaaatgtctcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 416
QY 483 gttgctcgccttgtagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 542
    |||
Db 417 gttgctcgccttcaataacatgctgag-----gtcccaactcgcgttcaactgtcacgga 470
QY 543 tcttctctcgcgcctccttgatggtgtggtcgtgaggaactaaccagctgagcgtac 602
    |||
Db 471 tcatctctcgtgcgccttgagcagatggtgcagagagacataccagcgcgtgacgcttc 530

```

QY	603	aaggattactatcccccaacgaagcagcgctgcggccgaatgcttggtaacatgacatgccc	662
Db	531	aaagactacttactcccccaataaagctacgctgcgtacccctatgtaaccaagatagct	590
QY	663	atgtccalccacccgcggaaacgcctccatctggtctcgaagctctgtctactatgtaccgaag	722
Db	591	atgcatactacgtctggaacgcctcaacgctgcgcagcgctgctctcaatgtctcaatgac	650
QY	723	ccggctctgagatctccctggaacctccccaagcgcgtacgcgcgagtttgatcccccttgct	782
Db	651	ccaaagccgaagagcgccttccaactctgtccaagctgtgatactgagatctgcataatgac	710
QY	783	ctgagctcccaagcgatataccaacgcagagagacacctctcttccacaacatgtgaagattcc	842
Db	711	ctcaagctccaagcagatatataccgcgaacgaagaaactgtgtacacacataatgtgagagctgac	770
QY	843	agcttctggggctgacgcttatctcaatgtgtaagtgtgagcccatgtgaatgcttcaagctc	902
Db	771	tcattcttgvggtgtgtatgaattccagct	796
QY	903	agaaatcatcgatgtatgaataattgtgcatgtctacagtgctataccaagaacggtacgc	962
Db	797	-----gaacggttcaac	807
QY	963	cttgcgctatgctcaacgctgcagccgcgcgaagtaaccgcgttccgcttccctcaacgctcgc	1022
Db	808	cgctgaccttcaagaagacgtgtgaagccctgcgaataatcatgattccgctctccctgcgtacgcag	867
QY	1023	ctctaacgctcttccgcctctctatctctgtacacctctgtgaagatctcaagaaacaaactccct	1082
Db	868	ttctcgcgtcttccgcgtcttactctgtcgtatctcgtatctcgtatcgaacatccgtcttgcctt	927
QY	1083	tccaggtcatctgcgcgtctgacagctgtgctctgcgtcttgagagccctgttgcacaactgtacactgct	1142
Db	928	tcaagttatattgctctcgcgtatctcgtatctctctcttgaacacacctgcgataccagcttgcgtt	987
QY	1143	acatctctatgcccgcgaagcgcgtctggaggtgtgtatctgcacttccacacttcgcgtcgcgt	1202
Db	988	acatttccatctgcgcgagcgcgttgaagaaattgtgtttgaacttccgcactatgctctgtgcgaaga	1047
QY	1203	ccatcgaatatactgcgaacctctccctctgctggtg-----ctgacgctctgcgtctgtgagccctgtgcttg	1259
Db	1048	ctattgaaacctgcgaacctctggcgtgtgatactgctgcgcgtacatgcgaaacagataccagctatg	1107
QY	1260	ataacactgtacaaaggtctatctgacattcgtctgtgataatgattcctctgtatgcgccgaacct	1319
Db	1108	acaacacacgcgaataaaggtctatctgcttggtagaagacagacaacatccgaacagataactc	1167
QY	1320	cttgaagctgacccgcgaacacctccgcgaagatgctctctccgcgaagcgacgaactctgagaccgcg	1379
Db	1168	cagttgttccctgcgtcaaaccttccgtgatgttccctctccctctcccaacacaaacaccccc	1227
QY	1380	caaaaccccaactgtacgagaaacttcaacctctgcgcgtgtctaaatgtgacagatgtgacatactca	1439
Db	1228	gaca-----gttccgcgttctgtctgcacgcgcgtctccactacatgtaactatca	1269
QY	1440	acggaagttaacctctcgtgatctgtagaagaaacggtctgtctccgaacatgtgcccgcgcgaactg	1499
Db	1270	atgtgtctgtcttctgtctgtatgttccaataacgcttgcgtctgtcaaacgttaccctgttgttatct	1399
QY	1500	cttgaagatctgcgcgaactgtgaataacaaactcccaacgcgtgttgagctaccctgttccatctaac	1599
Db	1330	ctggagctgttggagctatatacgaacgcgcgtgtgaacggttggagccgaacctatactcaatccatc	1389
QY	1560	tcgttgacttccgcgaatcttctctcgttc-----cagtcgccgttgagatctgagacctc	1610
Db	1390	ttgtcgcgaacttcaagtgtaattcttccgtacttccgcgaacaacagcgcgcacagatcatgcatat	1449
QY	1611	atgagagctgcgtgcgtctcaagaagatgtgtctgctgctgctctgctctgtatgagtgtgtctatggtt	1670
Db	1450	acagag-----tccggctctcaagaagacgttgcctgcgtctgtctgcgcgttgcgaataactgtgattgtt	1506

Accession	Sequence	Position
OY 1671	agggcactatagcgtccttccgtaagtctgcgtcttaactaatcgttttaactcat	1730
OY 1507	aggtcattaaagcgcctttccc-----	1528
OY 1721	gtcaacatctacaagctggtgtctacatglttgacatgcccacaactgataccagaggacca	1790
Db 1529	-----tgggtatcacatglttccatcttgccacaattgattatcaagaagatca	1573
OY 1721	gcacatgatgcgtctctcaatgatcaactcgtttcccggttgactatgtgctacaactaaccca	1850
Db 1574	cgaatgatgtgctgccttcaatgaacacgcgtctcctgcacagattgtgctataatgacacgt	1633
OY 1851	gttcaatgaccccatgtagcctctctctggaagcccgcctctcctctctcctcogaagattcga	1910
Db 1634	tttcgttgacacctatggaagaagctcttgtagcaggtctgtctcttcttgacatccgcagagttcca	1633
OY 1911	gaatgcgtcgcgtgtaactcgcagagcttgcacatctgacatgcacatctcggagatgtcgtg	1970
Db 1654	ggctcagagtgagcagatcgcgctctcaagctcgtttactatcgatcggtatctacagactatgkcgca	1753
OY 1971	cttcaacccttaagcccgagcgtatgta	1997
Db 1754	atacagaaccttaacgcgcgacgtatgaca	1780

RESULT	6
AAZ61243	
ID	AAZ61243 standard; DNA; 2110 BP.
XX	
AC	AAZ61243;
XX	
DT	30-MAY-2000 (first entry)
XX	
DE	DNA encoding a phenol oxidising enzyme.
KM	Phenol oxidizing enzyme; fungus; redox reaction; detergent; bleaching;
KW	fabric; pulp; paper; decolourisation; plant-derived food product;
KW	coloured compound; porphyrin; tannin; polyphenol; carotenoid;
KW	anthocyanin; Mallard reaction product; ss.
XX	
OS	Acremonium murorum.
XX	
Key	Location/Qualifiers
FH	135..1943
CDS	/*tag= a
FT	/product= "phenol oxidising enzyme"
FN	
XX	
PN	WO200005349-A1.
XX	
PD	03-FEB-2000.
XX	
PJ	13-JUL-1999; 99WO-EP04922.
PF	
PR	21-JUL-1998; 98EP-0202454.
XX	
PA	(UNIL ) UNILEVER NV.
PA	(UNIL ) UNILEVER PLC.
PA	(HIND-) HINDUSTAN LEVER LTD.
XX	
PI	Convents D, Gouka RJ, Van Der Heiden M, Swarthoff T, Verrips CT;
XX	
DR	WPt; 2000-195101/17.
DR	P-PSDB; AAY69204.
XX	
PT	Phenol-oxidizing enzyme from Acremonium, used in detergent compositions
PT	for bleaching stains on fabrics -
XX	
PS	Example 2; Page 37-40; 45pp; English.
XX	
CC	The present sequence encodes a phenol oxidizing enzyme from the
CC	fungus Acremonium murorum. The enzyme has the CBS accession number
CC	157.72. The enzyme catalyses redox reactions and is specific for
CC	molecular oxygen as the electron acceptor. The phenol oxidising enzyme

is specifically used in detergents for bleaching strains on fabrics, but also for bleaching pulp and paper and for decolourisation of plant-derived food products. The enzyme has a pH optimum in the alkaline to neutral range and can bleach a wide variety of coloured compounds, e.g. porphyrins, tannins, polyphenols, carotenoids, anthocyanins and Maillard reaction products.

Sequence 2110 BP; 400 A; 732 C; 556 G; 422 T; 0 other;

Query Match 24.7%; Score 517.4; DB 21; Length 2110;  
Best Local Similarity 61.4%; Pred. No. 1e-135;  
Matches 1041; Conservative 0; Mismatches 476; Indels 178; Gaps 6;

```

QY 321 accgtcccaacccaacactgtgagagagacattgttactacagagtgtgagatggccc 380
DB 390 accgtcccaacccaacactgtgagagagacattgttactacagagtgtgagatggccc 449
QY 381 ttctcccaacagatctacacgtgtctgagacgagcaacatgtgtgatatgagatg 440
DB 450 ttctctgagacaggtctctccctgacgtgagcctgagcctgtgtgtgtatgagcgc 509
QY 441 tcccaagacatacattcgtctcctgtgacactgtgagatgtgtgtccgttcgtgaa 500
DB 510 tcgcccgtctcctcattccacagagagagagagcgctgtcgtcttcgtgaa 569
QY 501 agcgagagagacacccctcccaacagcttccattgtcagcgctcttcttcagctccc 560
DB 570 a-----agccacgtctcgagagctccattccattccacgctcgttcagcgctccc 623
QY 561 ttctgattgtgtgagctgagacattaccacgctgtgagagatgacatctacc 620
DB 624 tgggaagcgctgagctgagatgtacacacccgagcgagacagactactatccc 683
QY 621 aacagcagagctgcccagatgtctgtgtacatgacatgacatgacatgacacgcgc 680
DB 684 aatgcgacagcgagcggtctcgtctgtacacatgacatgacatgacatgacacgcgc 743
QY 681 aagcgtacatgtgtgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 740
DB 744 aatgcgtacacagcgagcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 803
QY 741 aacgtcccaacgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 800
DB 804 ggtctgcacgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 863
QY 801 aacgacagagcgtctctctccacacatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 860
DB 864 aactcgcagctgtctctctccacacatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 923
QY 861 attcaagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 920
DB 924 atccaggt-----gaa 931
QY 921 aattgtgcatgtctaacagtgctatacaagaacgttcaagcgtgtgtgtgtgtgtgt 980
DB 932 -----gaa 980
QY 981 tgcagcgcgcgaagtcacgcttccgtctcccaacgctgtgtgtgtgtgtgtgtgtgt 1040
DB 1021 tctacttgtcgaacacgcgcgaagtcacgcttccgtctcccaacgctgtgtgtgtgt 1080
QY 1041 tgtacttcttactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1100
DB 1101 acggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1160
DB 1081 acctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
QY 1161 gctgagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1220
DB 1141 gctacagatcattcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1280

```

```

QY 1221 ttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1280
DB 1201 agccgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1260
QY 1281 gatt---cgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1337
DB 1261 gcttcaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1320
QY 1338 tccgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1397
DB 1321 tccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1365
QY 1398 agacttcaacttccgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1457
DB 1366 actcgttccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1425
QY 1458 atgtcgaagacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1517
DB 1426 actcgtgaagacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1485
QY 1518 agaacacttcaacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1577
DB 1486 cgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1545
QY 1578 ttctc-----gttccacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1619
DB 1546 tgtctgcacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1605
QY 1620 ctgtctcgaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1679
DB 1606 ctgtctcgaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1665
QY 1680 acgctccttccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1739
DB 1666 acgctccttccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1679
QY 1740 tacaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1799
DB 1680 -----ggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1732
QY 1800 ggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1859
DB 1733 ggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1792
QY 1860 ccccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1919
DB 1793 ccccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1852
QY 1920 ggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1979
DB 1853 gaacgccttcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1912
QY 1980 ctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1994
DB 1913 ctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1927

```

# RESULT 7

AA227601  
ID AA227601 standard; DNA: 1791 BP.

AA227601:

16-DEC-1999 (first entry)

Stachybotrys phenol oxidase coding sequence.

Phenol oxidase; enzyme; coloured compound; dye transfer prevention;

fabric washing; stain bleaching; anti-dye transfer; detergent; ss.

Stachybotrys chartarum.











```

Db 715 aacgcgaatgtaaccctcgctcgaccgaggtgagaccaggaacctggtggagatgct 774
QY attcaaatgtaagttagaccatctgagatctcaagatcagaatgatactgtatga 920
Db attcaatgctc----- 783
QY 921 aattgcatgctctaaccaagtgctatcacagaacgtacagcttggtcctacagc 980
Db 784 -----aaagacagacatgagcttctcctaag 811
QY 981 tgcagcgcgcgaagtacgccttcgccttcctcaagctgcgcctcaagcttctgcctc 1040
Db 812 tccagcccccgaatcagcttctcgaatctcctcaagctgcgcctgcgtcttgctcc 871
QY 1041 tgcatactgtaacctctgaagatctcagagacagcttccctcagagtgatctgcgcgt 1100
Db 872 tctaacctgcgcaggaacagctctcccaagctcagaatctctcctaagtgatctgcctctg 931
QY 1101 acggtgctgctgcttgagggccctgcttgacacatctgacatctctacatctctaaagccgagc 1160
Db 932 atgctcgtctccttcaagcccccgctcaagacctcaacctctgctgctgctgcgagc 991
QY 1161 gctggagaggtgtatcagacttctcaacctctgctgcccagctcatalcagatccgcaac 1220
Db 992 gttaacagacatctatgacttcaaccaactctgctgcagactcttgacctgcgcaacg 1051
QY 1221 ttcctgtgctgacaggtctcgcgtgtgagcctgagtgtgataacacatgacaagtgatgc 1280
Db 1052 ttgctgaagacaacagatgtcgtgcagagatgagatagctgcacactctcgaagtgatgc 1111
QY 1281 gattcgtcgttgatgagatctcctgagtcgcgcgaacctctgaggtgctgccaacctc 1340
Db 1112 gctcgtcgtcagctcgtgacgtgtgag--gacaacagcagctcctcctcaactctcc 1168
QY 1341 gaagatgctcttcccccgaagggcgcaactgtagagcccccgaacccagatgacagaga 1400
Db 1169 gtgagcttctctcc-----tcccaagaagaagccccgcgcaacagc 1213
QY 1401 cttcacctctcgagcgtgtaataagacagtgaacaatcaacagagatctctcgatg 1460
Db 1214 actcaagtttgagacgagacagacgacatactgataacagatgtgtgttcgcagatg 1273
QY 1461 tgcagaaacgctcgtcgcgaatgtgccccgcgacactgtgtgatacttgagcagcttgaga 1520
Db 1274 tcaataagcgtctcgtccgcaaacgcaagctgcgcagcgtgtgagcttgagagctcgaga 1333
QY 1521 acaacaccaaagcttggaacacacccgttcaacatcaacccgttgatctccgagctctt 1580
Db 1334 acccctcgtgagcgtgagacacccgcgtcaacatctcaactgtgtgatacctaagatctca 1393
QY 1581 ctgcttcaactgcgcgtg--agtcgagcctatgagcgtgctgtcctcaagatgtg 1637
Db 1394 agcgaactgtgtgctggtgcgacagctgcctcaagctgtgctgtcttaagatgtcg 1453
QY 1638 tctggtgctgctgctgctgaggtgtgtatgtgtgagcccaactacgctcttccgtaag 1697
Db 1454 tctggttggcaggggtgagacccctgacatcagagcccaactcaacaccttgac----- 1508
QY 1698 tctcgccttcaactaactgcttcaactcaatgatacatacacaagtggttctacat 1757
Db 1509 -----tggaacttaacat 1520
QY 1758 gttgcaactgcacaacatgatacagagagacagacatgatagtgcttctcaatgtaac 1817
Db 1521 gtgacactgttcaacatcatcttaacagagatacagatgatagtcgttattcaacgtcac 1580
QY 1818 tgttctcgtgatagtgtacacatacagacagctgctatctgacccaatgtagctctcgt 1877
Db 1581 cgcacatgagagaaagagatatcttaagagagaccttcgagagcccatgagaccccaagtg 1640
QY 1878 gaagcccccgcctctcctcctcgagagatcgagaatgagctggtcggtacagagacgagc 1937
Db 1641 ggcgcgcgttctcttaaacaccacacgacttctcgtcgcgtcggaactctccgcgga 1700

```

```

QY 1938 tgcatactgacgcacatcagagatggtgacttcaacccctacgccccaggtgatga 1997
Db 1701 gtccatcacctgcgcgcaggtgagagagctggtgcgagcagagcgatcaacacgctgatga 1760

```

```

RESULT 11
AA25735
ID AA25735 standard; DNA; 2067 BP.
XX
AC AA25735;
XX
DT 05-JAN-2000 (first entry)
XX
DE Stachybotrys chartarum phenol oxidising enzyme PCR fragment.
XX
KW Stachybotrys chartarum; phenol oxidising enzyme; colour; dye;
KW detergent; anti-dye transfer; stain removal; bleaching; PCR primer; ss.
XX
OS Synthetic.
XX
SS Stachybotrys chartarum.
XX
PN MO9949010-A2.
XX
PD 30-SEP-1999.
XX
PF 23-MAR-1999; 99MO-EP02042.
XX
PR 24-MAR-1998; 98US-0046969.
PR 22-DEC-1998; 98US-0218702.
XX
PA (UNITL ) UNILEVER NV.
PA (UNITL ) UNILEVER PLC.
XX
PI Convents D, Amory A, Wang H, Dhaese P, Lambrechts-Rongvaux A;
PI Wang C;
XX
DR WPI: 1999-601211/51.
XX
PT Detergent composition containing phenol oxidase from Stachybotrys, used
PT to bleach stains and prevent dye transfer
XX
PS Example 17; Fig 9; 56bp; English.
XX
CC The present invention describes a detergent composition containing a
CC purified phenol oxidising enzyme derived from Stachybotrys. The present
CC sequence represents a PCR fragment of Stachybotrys chartarum phenol
CC oxidising enzyme. The enzyme can be used to modify the colour of dyes
CC and other coloured compounds (e.g. for use in pulp and paper bleaching
CC also for removing stains, e.g. food, tea, blood etc., from fabrics) and
CC for preventing dye transfer during fabric washing.
XX
SQ Sequence 2067 BP; 445 A; 621 C; 500 G; 501 T; 0 other;

```

```

Query Match 14.3%; Score 300.4; DB 20; Length 2067;
Best Local Similarity 55.2%; Pred. No. 2,4e-74;
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;
QY 380 ctctccacacgaatctacactgatactgagccgcgcacaatggttgatagatgact 439
Db 457 ctgaccccttcagatttaaccacaccccttgccctgcacactcgtcgtgacagatgcat 516
QY 440 gtcccagagacacatcatcgtctcgttgagcactgagaggtgtcgcgtctgga 499
Db 517 gagcctgtctcacttcaatgattctccagagagaaagagactggtatgattacaa 576
QY 500 cagcgagagaaacactctccacagcgtccacttgacagcgtcttctctcagactcc 559
Db 577 ca-----atgcacacgttgagacatcgtlccatctgcacagctcccatcggtcccc 630
QY 560 ctttgatgtgttggtcgagacactacccagctgtgcaggtataaagattactacacc 619

```

Db 631 ttccgatgtgttggtcggaagatgtgaccttccttcgtcgagttacaagattactattcc 690  
 QY 620 caacaaagcaggtcgccgcagatgtcttggtacataagacatgc----- 661  
 Db 691 caactaccacatccgcgcctctctgtgtacacatgacacagcttcatgaagtatgcta 730  
 QY 662 -----catgtccacacgcgcgag 680  
 Db 751 cgagccttactcttctgtctactcttggctaaccaacttcttctgtagtctgtgag 810  
 QY 681 aagccttaactgtgtgaaggtgtgtctaatatataccagaccgcgtgaggtgcctcg 740  
 Db 811 aatgctactctgtgtcaggtgtgcctacatatatacaacgacgaggtgaggaatgcttc 870  
 QY 741 aacctccacgagcgacgacgaggttgatatacccttctgtcttgagctgcacagcatac 800  
 Db 871 ggtcttcctcctatgtgtcatgtgaggttcgatatatccctctgtacttgacgacgaat 930  
 QY 801 aacgcagacggaactctctctccacaaatgagaggtttccagctcttgaggatgagctt 860  
 Db 931 aacgcagatgtacactctgtcgcacgaggggtgaggaacaggaactctgtggagatgtc 990  
 QY 861 attcaaggtgtgaagttgagcccatitgagatgtctcaagatccatgaagatcatgata 920  
 Db 991 atccatgtc----- 999  
 QY 921 aattgtcatgtctataacagatgtataccagacgctgcagcttgctatgtctaacg 980  
 Db 1000 -----aacggaacgacatgtcttcttcttaacg 1027  
 QY 981 tgcagcgcgcgaagtaacgccttcgccttccaaacgctgcgtctcaacgcttctgcctc 1040  
 Db 1028 tccagcccgccgaagtaacgcttctcgaattcccaacgctgcgtctgcgtctgcctcc 1087  
 QY 1041 tgaattcttactctgtgaggtatcagagaccgacttccctccaggtcatgtgcgcgtcg 1100  
 Db 1088 tcaactctgtcaggaacagctccccaacgttcagaaatctcttccaagtcatgtgctctcg 1147  
 QY 1101 acggtgtctgtcttgaggccctgtgtgaacactgtacactctgtacatctatgagccgagc 1160  
 Db 1148 atgctgtctctcttcaagcccgcttcagacactctacactctgtctgttcgcgcagc 1207  
 QY 1161 gctggagaggtgtta----- 1175  
 Db 1208 gttacgagatcatattgtgtatgacctccctccctccacgaatgagtaagaactctaagac 1267  
 QY 1176 -----tgaacttccacctctgcgtgcgcagctcatcagatccgaacttccctcg 1226  
 Db 1268 taacactgttagacttccacaaacttgcgtgcagactcttgacctgcgcgaactgtgctg 1327  
 QY 1227 gtgctgacgctcgtgtgtgagcgtgatttgataacactgtgacaggtcatgtcgattcg 1286  
 Db 1328 agaccaaagatgtcgcgcagagatgagatgcgcacacccgcagagtgagatgcgcctcg 1387  
 QY 1287 tgcgttgaatgaatctcctgtgcgcgcgaacacttctgaggtgccttccaaactccgaagt 1346  
 Db 1388 tgcgtacgctgtgcactgtgtgag---gacaacagcaggttccctccactctccgtgagcg 1444  
 QY 1347 ttcccttcccgagggcggaactgtgaccccgcaaaccccaactgtatgagagactctga 1406  
 Db 1445 ttccttctcct-----cctacaagaagaagcccgccgcagacaagacttca 1489  
 QY 1407 ccttcgacgtgtctaattgacagttgacacataacgaggttactctctcgatgtgagaga 1466  
 Db 1490 agtttgaacgcagcagcagacactactatcaacgattgtgcttgcgcgagatcaatg 1549  
 QY 1467 accgttctgcgcaatgtgcgcgcgacactgtgtgagatcttggcagactgtgaacaact 1526  
 Db 1550 agcgttctctgcgcgaacccgcagagctgcgcacgcttgaggtcttgaggcctccgaactcct 1609  
 QY 1527 ccaaggttgagactaacctctgtcaacatacactcgttgaactcgcagcttcttccgtt 1586  
 Db 1610 ctggaggtcgtgagccacccgctccacatcactctgttgaactcaagatccctcaagcgaa 1669

QY 1587 ccactgcgcgttg---agtcagccttatgaggtcgtgtctccaaagatgtctctgcgc 1643  
 Db 1670 ctgtgtgtgtgtgcgcaggtatcatgtcctcctacagatctgtctgtcttaagatgtctgtc 1729  
 QY 1644 tggcctgcgtgtgaggtgtgtctatgtgtgagggcccaactaagctccttccgttaagtctcg 1703  
 Db 1730 tgggcaggggtgagagaccttgaccatcgagggcccaactaaccacttgac----- 1778  
 QY 1704 cctttacactatgttttcaactatgtcataactacaaagtggtgtcacaatgttgcga 1763  
 Db 1779 -----tgagacttaacatgtgcga 1796  
 QY 1764 ctgcacaacctgatccacgaggaaccaacagatgatgtgtcttcaatgaacttctc 1823  
 Db 1797 ctgtcaacaactcatccacagagataacagacatgagctgtatccaacgtccacat 1856  
 QY 1824 cgggtgactatgtctacaactacacagatgtcatltgaccccatggaagctctctgagagcc 1883  
 Db 1857 ggaaggaaggaatatcttcaagagagacttcgagagaccatgaaccccaagtgcgcgcg 1916  
 QY 1884 cgcgcctctctctcgtgagaggttcgagaaatgtctcgggtgacttcaagagcttccat 1943  
 Db 1917 cgttcttcaaacgcgaacagacttccatgtcgcgtggaacttctccgcgcagatccat 1976  
 QY 1944 cacttgcagatcagagagatgtgcttcaaacccctacgcagagctgtatga 1997  
 Db 1977 cactgcgcagatgtcagagagctgtgcgcgacgagaccgttaaacgcctcgatga 2030

RESULT 12  
 AA227609  
 ID AA227609 standard; DNA; 2067 BP.  
 XX  
 AC AA227609;  
 XX  
 DT 16-DEC-1999 (first entry)  
 XX  
 DE Stachybotrys phenol oxidase coding sequence.  
 XX  
 KW Phenol oxidase; enzyme; coloured compound; dye transfer prevention;  
 KW fabric washing; stain bleaching; anti-dye transfer; detergent; ss.  
 XX  
 OS Stachybotrys chartarum.  
 XX  
 PN WO949020-A2.  
 XX  
 PD 30-SEP-1999.  
 XX  
 PF 23-MAR-1999; 99WO-US06327.  
 XX  
 PR 24-MAR-1998; 98US-0046969.  
 PR 22-DEC-1998; 98US-0218702.  
 PR 22-MAR-1999; 99US-0273957.  
 PA  
 PA (GENEW) GENENCOR INT INC.  
 XX  
 PI Amory A, Wang H, Dhase P, Lambrechts-Rongvaux A, Wang C;  
 XX  
 DR WPI; 1999-591088/50.  
 XX  
 PT Novel enzyme for modifying coloured compounds used to prevent  
 PT dye-transfer -  
 XX  
 PS Example 17; Fig 9; 64pp; English.  
 XX  
 CC This sequence encodes the Stachybotrys chartarum phenol oxidase enzyme  
 CC of the invention. The invention is used to modify a coloured compound and  
 CC prevent dye transfer during fabric washing, or for stain bleaching or  
 CC anti-dye transfer. It is useful in the detergent, paper and pulp, textile  
 CC and food industries.  
 XX  
 SQ Sequence 2067 BP; 445 A; 621 C; 500 G; 501 T; 0 other;

Query Match 14.3%: Score 300.4; DB 20; Length 2067;  
 Best Local Similarity 55.2%: Pred. No. 2,4e-74;  
 Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

```

QY 380 ctctcccaacatctaccctgcatctgagccgccaacatggttgatagatggcat 439
    || || || || || || || || || || || || || || || || || || || ||
DB 457 ctgacctcttcagattaccctccaccttgccctgcgcactctgcgcgaatggcat 516
    || || || || || || || || || || || || || || || || || || || ||
QY 440 gtccccaagaccatcatatctgctctgctgagactgagagtggttcgcctgtgaa 499
    || || || || || || || || || || || || || || || || || || || ||
DB 517 gagcccttgctccacttcaatgctcccaagagacaagactgtaagttcaatcaa 576
    || || || || || || || || || || || || || || || || || || || ||
QY 500 cagcgagagaaacactctcccaacagcgtctccactgtcaagcgctcttctctgaatcc 559
    || || || || || || || || || || || || || || || || || || || ||
DB 577 ca-----atgcacccgttgagaaactcgtctccatcgtcaagcgctcccaatcggtgccc 630
    || || || || || || || || || || || || || || || || || || || ||
QY 560 ctltgatgttggtgtgaggaacataccagcctggcgagttacaagattactacc 619
    || || || || || || || || || || || || || || || || || || || ||
DB 631 ttccgaatgttggtgtgaggaatgtgaccttccctggcgagttacaagattactacc 690
    || || || || || || || || || || || || || || || || || || || ||
QY 620 caacagagcagcgctgcccgaatgtgttgatccatgacatgc----- 661
    || || || || || || || || || || || || || || || || || || || ||
DB 691 caactacaatccgcgcgcctctctgtgtacatacgaacagccttcaatgaagttatgta 750
    || || || || || || || || || || || || || || || || || || || ||
QY 662 -----catgtccatcacgcgcgag 680
    || || || || || || || || || || || || || || || || || || || ||
DB 751 cgagccttattcttctgtgtacacttggtaacaaactcttccgtatagatgtgag 810
    || || || || || || || || || || || || || || || || || || || ||
QY 681 aacgcctacatggtgtagcgtgtgtctacatgataccaaagaccgcgtgagatgcctg 740
    || || || || || || || || || || || || || || || || || || || ||
DB 811 aatgcctactctgtgtgtagcgtgtgtgcctacatatacaacgagcgtgagatgtctc 870
    || || || || || || || || || || || || || || || || || || || ||
QY 741 aacctcccaagcgcgtgacgaggttgatctcacttgcctgtgtgtgctgcgcgagatac 800
    || || || || || || || || || || || || || || || || || || || ||
DB 871 ggtcttccctgaagtgatagtgagagtgatgataccctcctatccctgcgcgcaagatatac 930
    || || || || || || || || || || || || || || || || || || || ||
QY 801 aacgcagacgagcactctctctccaaatgagaggttccagcttctgtggtgagat 860
    || || || || || || || || || || || || || || || || || || || ||
DB 931 aacgcagatgtatccctgtgtctgacccgaggtgagagacagagacgtgtggtgagatgtc 990
    || || || || || || || || || || || || || || || || || || || ||
QY 861 attcaagtgtgtaagtgtgagccatgtgagatgtctcagatccctagaagtatcagatgta 920
    || || || || || || || || || || || || || || || || || || || ||
DB 991 attcaatgtc----- 999
    || || || || || || || || || || || || || || || || || || || ||
QY 921 aattgtgcatgtcttaaccagttgcatcaagaacggtcaagccttgcatatgctcaacg 980
    || || || || || || || || || || || || || || || || || || || ||
DB 1000 -----aacggaacagcaatgtgccttcttcaacg 1027
    || || || || || || || || || || || || || || || || || || || ||
QY 981 tgcagcgcgcgaagtaacgcttccgcttccctcaacgctgcgtctcaacgcttctgcctc 1040
    || || || || || || || || || || || || || || || || || || || ||
DB 1028 tccagcccccgaagtaacgcttccgattctcctcaacgctgcgtgcgtctgcgtctc 1087
    || || || || || || || || || || || || || || || || || || || ||
QY 1041 tgtatctgtctactctgtgagatctcaagaacagacttccctcccaagttatgcgcgtg 1100
    || || || || || || || || || || || || || || || || || || || ||
DB 1088 tctactctgcagagcaacgtctcccaacggtcagaattcttcccaagtatgtcctctg 1147
    || || || || || || || || || || || || || || || || || || || ||
QY 1101 acggtgtgtgtctgtgagggcctgtgtgacactgtgatacactgtacatgtcagcgacg 1160
    || || || || || || || || || || || || || || || || || || || ||
DB 1148 atgctgtgtctctctcaagcccgcttcagactcttaactcttaactgtgtgtgcgagag 1207
    || || || || || || || || || || || || || || || || || || || ||
QY 1161 gctgggaggttgtta----- 1175
    || || || || || || || || || || || || || || || || || || || ||
DB 1208 gttaacgaatcatattgtgtatgacctccctccatcaagatgtgtaagaactctaagac 1267
    || || || || || || || || || || || || || || || || || || || ||
QY 1176 -----tgactctccactctgcgtgtgacatgcatacgatgcgaacttccgtg 1226
    || || || || || || || || || || || || || || || || || || || ||
DB 1268 taacactgttagacttcaacaaacttgcgtgcagacactcttgacccgtgcgaagttgctg 1327
    || || || || || || || || || || || || || || || || || || || ||
QY 1227 gtgctgtagcgtctcgtgtgtgagcgtgagttgatacactgtgcaaggtcatgctgctg 1286
    || || || || || || || || || || || || || || || || || || || ||
DB 1328 agagcaacagatgtcgtgcagagagatgatacgtctgcacactctcaggtgtgagtgccttg 1387
    || || || || || || || || || || || || || || || || || || || ||
  
```

```

QY 1287 tctgtatgaagtccttgagtcgcccgaacactctgagtgtcctgcgaactccgagatg 1346
    || || || || || || || || || || || || || || || || || || || ||
DB 1388 tctgtacgtctgtgcactgtgtgag-----gacaacagccaggttccctccactctcgagag 1444
    || || || || || || || || || || || || || || || || || || || ||
QY 1347 ttccttccccgaaggcggaactggaacccgcaaacccactgtatgacgaacttca 1406
    || || || || || || || || || || || || || || || || || || || ||
DB 1445 ttccttctccct-----cctcaagaagaagccccgcgcgcgaacgaacttca 1489
    || || || || || || || || || || || || || || || || || || || ||
QY 1407 ccttcggcgctgtctaattgagcagtggaacaaactcaagagagttactctctgtgagtggaga 1466
    || || || || || || || || || || || || || || || || || || || ||
DB 1490 agtttgaacgcagacagacactactactgatacagatgtgtgcttgcgatagtcaatg 1549
    || || || || || || || || || || || || || || || || || || || ||
QY 1467 accgtctgtctccgaatgtgtgccccgcgaacactgtttagatctgtgcagatgaaacaaact 1526
    || || || || || || || || || || || || || || || || || || || ||
DB 1550 agcgtgtctcgtgcgaagcccgagctgcgaacgcttgaggtctgtggagctgcgaactcct 1609
    || || || || || || || || || || || || || || || || || || || ||
QY 1527 ccaacggttgagactcaacccgtgtcaatcaactcgtgtgaactccgagttccttctglt 1586
    || || || || || || || || || || || || || || || || || || || ||
DB 1610 ctggaaggttgagacacaccccgctccacatccaactgttgaactcaagatctcaagcgaa 1669
    || || || || || || || || || || || || || || || || || || || ||
QY 1587 ccaactgcccgttg-----agtcgagccttatgaggtctgtgtcgaagagttgtctgac 1643
    || || || || || || || || || || || || || || || || || || || ||
DB 1670 ctggtgtgtcgttgagccaggttcaatgcccctacagatctgtgtcttaagagatgtctgtg 1729
    || || || || || || || || || || || || || || || || || || || ||
QY 1644 tggctcgtctgaggtgtgtctatgttgaagcccaactacgctcttccgtaagttctcg 1703
    || || || || || || || || || || || || || || || || || || || ||
DB 1730 tggcagaggtgtgagacccctgacatctgagggccacatacccaaccccttgac----- 1778
    || || || || || || || || || || || || || || || || || || || ||
QY 1704 cctttaactaactgttcttaactcatgtcaatctaacatgaagtgtgtctacatgttga 1763
    || || || || || || || || || || || || || || || || || || || ||
DB 1779 -----tgagagttactatgtgta 1796
    || || || || || || || || || || || || || || || || || || || ||
QY 1764 ctgcacaacactgataccaagagacacagatgagatgtgtgcttcaatgtacttctc 1823
    || || || || || || || || || || || || || || || || || || || ||
DB 1797 ctgtccaaactcatccaagagataacagatgagatgtgtgttcttaacgtccgcacat 1856
    || || || || || || || || || || || || || || || || || || || ||
QY 1824 cgtgtacatgtgtacaaactacacagagttcatctgaaccccaatgagagctctctggaagcc 1883
    || || || || || || || || || || || || || || || || || || || ||
DB 1857 ggaaggaagaggaatatcttaagagagacttcgagagcccatgaaccccaatgagcgagc 1916
    || || || || || || || || || || || || || || || || || || || ||
QY 1884 cgcgccttccctctcgtgagaggttcgaagatgtcgtgtgagttcaagagagttgccat 1943
    || || || || || || || || || || || || || || || || || || || ||
DB 1917 cgttcttaaacacgcaacgaacttcaatgtctgcgttggaataacttctcgcgcgagttccat 1976
    || || || || || || || || || || || || || || || || || || || ||
QY 1944 cactgagccatccaagagatgtgacttcaacccctcagcccaagcgtatga 1997
    || || || || || || || || || || || || || || || || || || || ||
DB 1977 cactgcccaggttcagagagctggtgcgagcagagagccgtataaacgcctcgatga 2030
    || || || || || || || || || || || || || || || || || || || ||
  
```

## RESULT 13

AA25728  
 ID AA25728 standard; DNA; 3676 BP.

XX AA25728;

AC 05-JAN-2000 (first entry)

XX Stachybotrys chartarum phenol oxidising enzyme genomic DNA.

DE Stachybotrys chartarum; phenol oxidising enzyme; colour; dye;  
 KW detergent; anti-dye transfer; stain removal; bleaching; ds..  
 XX  
 OS Stachybotrys chartarum.

PN WO949010-A2.

XX 30-SEP-1999.

PF 23-MAR-1999; 99WO-EP02042.

XX 24-MAR-1998; 98US-0046969.  
 PR 22-DEC-1998; 98US-0218702.



RESULT 14  
AAZ27602  
ID AAZ27602 standard; DNA; 3677 BP.  
XX  
AC AAZ27602;  
XX  
DT 16-DEC-1999 (first entry)  
XX  
DE Stachybotrys phenol oxidase genomic sequence.  
XX  
KM Phenol oxidase; enzyme; coloured compound; dye transfer prevention;  
XX fabric washing; stain bleaching; anti-dye transfer; detergent; ss.  
OS Stachybotrys chartarum.  
XX  
PN WO9949020-A2.  
XX  
PD 30-SEP-1999.  
XX  
PF 23-MAR-1999; 99WO-US06327.  
XX  
PR 24-MAR-1998; 98US-0046969.  
XX 22-DEC-1998; 98US-0218702.  
XX 22-MAR-1999; 99US-0273957.  
PA (GENEV ) GENENCOR INT INC.  
XX  
PI Amory A, Wang H, Dhase P, Lambrechts-Rongvaux A, Wang C;  
DR WPI; 1999-591088/50.  
DR P-PSDB; AAY39992.  
XX  
PT Novel enzyme for modifying coloured compounds used to prevent  
XX dye-transfer.  
XX  
PS Claim 21; Fig 6; 64pp; English.  
XX  
CC This sequence encodes the Stachybotrys chartarum phenol oxidase enzyme  
CC of the invention. The invention is used to modify a coloured compound and  
CC prevent dye transfer during fabric washing, or for stain bleaching or  
CC anti-dye transfer. It is useful in the detergent, paper and pulp, textile  
CC and food industries.  
XX  
SQ Sequence 3677 BP; 822 A; 1056 C; 849 G; 947 T; 3 other;

Query Match 14.3%; Score 300.4; DB 20; Length 3677;  
Best Local Similarity 55.2%; Pred. No. 3.3e-74;  
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

QY 380 cttctccaccagatctacatctgtgagcgcgccaacatgtgtgataagatgcat 439  
DB 1491 cgcacccttcagatctacccacacttcgcgcctgcaactctgcgtacagatggcat 1550  
QY 440 gtcccaagaccatcatcatcgttcctcgtgacatgagatgtgtccgcgttgaa 499  
DB 1551 gaggcctgtctcacttcaatgttccagagaggaagacagctgtatgtatcga 1610  
QY 500 cagcgagagaaacactctcccaacagcgtccaactgcaagcgtcttctcagctcc 559  
DB 1611 ca-----atgcacacgttgagaaactcgttcacatcgtccacatcggtgcc 1664  
QY 560 ctttagtggttgagtgagagacacacacacgctgagagataagattactaaccc 619  
DB 1665 ttcgtatgtgtggcgtgaagatgtgacctccctcgtcgagataagattactacttcc 1724  
QY 620 caacaagcagcgtccgcagatcgttgatacattgaccatgc-----661  
DB 1725 caactacacatccgcgcgtctcgtgtacacatgacacagcgtcttaataagtatgcta 1784  
QY 662 -----catgtccatcacgcgcag 680

DB 1785 cgagccttactcttcttgtaactcttgtaaccaacttctctgtagactgtgag 1844  
QY 681 aacgcctacatgtgtacagcgtgtgtctacatgatccagacccggttagatgccc 740  
DB 1845 aatgcctactctgttgaagcgttgccatcatatataacagagcgtgagatgtctcc 1904  
QY 741 aactcccccagcgcgtacagcagattgatalcccttgcttgactgccaagcatac 800  
DB 1905 gttcttccatgagtgcatgtgagattcgaatccctgatccctgacgcgccaagtactat 1964  
QY 801 aacgcagcgcacatctctctccaccaatgagagtttccagcttcgtggttagcgt 860  
DB 1965 aacgcagatgtaacctgcttgacacagaggttgagacacagactgtgggagatgtc 2024  
QY 861 attcaagtgttaagttgagccattgagatgcttccagatccatagaagtatcgatgta 920  
DB 2025 attcatgtc-----2033  
QY 921 aattgtcatgtcttaacacagatgctatcacagacggtcagcttgctatgccaag 980  
DB 2034 -----aacgcagacgatgtgcttcccttaacg 2061  
QY 981 tgcagcgcgcgaagtaacgcttcgccttccatcaagctgcgtctcaagctcttgcctc 1040  
DB 2062 tccagcccccgaagtaacgcttccgatcttccatcaagcgtgcgtctcgtctgtcctcc 2121  
QY 1041 tgtatcttgctacacttgaggttcagagacagacttcccttcacagatgctgcgctg 1100  
DB 2122 tctaacctgtcagagacagctccccaacgtcagaatcttcttccatgaatctgctctg 2181  
QY 1101 acggtgtctgttgagggccctgtgtacacatgacatctgtacatctatagccagag 1160  
DB 2182 atgctgtctccttcaagcccgctcagacacttaaccttaacctgtgtgtccgagc 2241  
QY 1161 gctggaaggtgtta-----1175  
DB 2242 gttacgagatcatattgtgtatgcctccctccctccatcagaatgataaagactaagac 2301  
QY 1176 -----tcgactctcacctcgtcgtgcgcagatcagatcagatccgaactctccg 1226  
DB 2302 taacacttgtatgactacacaacttgcgtgcagacgtcttgacgcgcgaagttgtcg 2361  
QY 1227 gtgtgacggtcgtcgtgtgttgagctgagttgataacactgacagatcagatcgatcc 1286  
DB 2362 agaccaagatgtcgcgcagcagatgagtagctgcgaactcgtgaggtgagtgcttcg 2421  
QY 1287 tggttatgaagtctctgtgagtcgcgcgacactcttgaagttgctgcgaactccgagatg 1346  
DB 2422 tcttcagctctgtgcaactgttgag--gacaacagccaggtlccctcactctcgttgacg 2478  
QY 1347 ttccttcccgagggcgcaactgtggaccgcgcaaaccccaactgtatgacgagacttca 1406  
DB 2479 ttccttccct-----cttcaagaagaagccccgcgcgaagaacttca 2523  
QY 1407 ccttcgcgcgttaatgacagtgagacaatacagaagatataccttctcgtatgtcagaa 1466  
DB 2524 agtttgaacgcagcaagacgaactactgtatcaagatgtgtgcttgcgtatgtcaatg 2583  
QY 1467 accgtctgtctcgcgaatgtgccccgcgcgacacgttgtagatcgtgagcattgagaact 1526  
DB 2584 agcgtgtctctgtgcgaagcccgagatctgacacgttttgaagctcgtgagacatgactct 2643  
QY 1527 ccaaggttgtagctacacccctgttcatcatcactcgttgacttcgcagctccttcgtt 1586  
DB 2644 ctggagagctggagacaccccgcttcaatcaactctgttacttaagaatcctccaagcgaa 2703  
QY 1587 ccaactgccctgtg--agtcgaagccttlatagagctgtgtctcaaggaatgtgtctcggc 1643  
DB 2704 ctgtgtgtgtgtgccaagfcatgtccctacagagctgtcgtcttaagagatgtctgtgt 2763  
QY 1644 tggctcgtctgtgaggtgtgtctatgtatgtgagcccaactaagctccttcccgtaagttctcg 1703  
DB 2764 tgggcaggggtgagacccctgacatcagagcgccactaacaaccccttgac-----2812



Db 1965 aagccgagtgtaacccctcgttcgagccgagggtagagaccaggaacctggtggagatgttc 2024  
 Qy 861 attcaagtgttaagtgtgagcccatgtgagtgtctacagctccagaagtatcatgtatba 920  
 Db 2025 atcaatgtc----- 2033  
 Qy 921 aatgtcatgtcttaacacagltgtatacagaagcgttaagccttggcctatgtccaag 980  
 Db 2034 -----aacgagacgcatgtccttccaaag 2061  
 Qy 981 tgcagccgagcaagtaacgcttcgccttccttaacgcttgccgttcctacgcttcttcgcctc 1040  
 Db 2062 tccagcccgagcaagtaacgcttctcgaattcctcaacgctgcgttcgtctgttcgtcc 2121  
 Qy 1041 tgaatcttgatccctcctgtgagatcagagatccagacttccctccaggtcatgtccgctg 1100  
 Db 2122 tctacccctgcagagacagctctccaaagctagaatctccttccaaagtcatgtccttg 2181  
 Qy 1101 acggtgtctgtcttgagggccctgttgacaactgtacaactgtacalctcatgagccgagc 1160  
 Db 2182 atgctgtctctcttaagcccgcttcagaccttaaccttaacctgtgtgttcgagagc 2241  
 Qy 1161 gctggaggtgtgtta----- 1175  
 Db 2242 gttaagagatcatltagtgatccctccctcctcaagaatgagtaagaactctaagac 2301  
 Qy 1176 -----tgacttctccactcgtgcagcttcgacttcgatatccgcaaccttcgt 1226  
 Db 2302 taacactgttaagacttcaacacttctgcgcagactcttgactctgcgcagagcttgctg 2361  
 Qy 1227 gtcgtgaagctcgtgtgttgagccgtgagttgataacactgaacaggtcatgtgctg 1286  
 Db 2362 agaccaagatgtcgcgcagagatgagtaagctgcgactctgcgaggtgagcttcgc 2421  
 Qy 1287 tgcgtgaagatgctcttgaggtgcgcggaactctgaggtgtcctgcgaactccgagatg 1346  
 Db 2422 tgcgtgaagctcgtcagctgttgag--gacaagagccaggtccctcactcctcgttgagc 2478  
 Qy 1347 ttccttcccgagggcggaactgtgagcccgcaaccccaactgaatgaagagacttca 1406  
 Db 2479 ttccttccct-----cctcaagaagagcccgcgagcaagcaacttca 2523  
 Qy 1407 ccttcgagccgtgtctaattgagacagltgaaacatcaacggagttacttctcgatgtcgag 1466  
 Db 2524 agttgaagcgagcaacgacactaactgtatcaacagatgtgtgttcgagtgatga 2583  
 Qy 1467 accgtctgtctcgcgaatgtgcccgcggaactgttgagatctgcgcgacttgagaacaa 1526  
 Db 2584 agcgtgtcctcgcgaagcccgagactcgtgaggtcgtgaggtcgtgaggtcgaagaactc 2643  
 Qy 1527 ccaacggttgagactaacctgttcaacttcaactcgttgaacttcgaagtccttctcgt 1586  
 Db 2644 ctggaggtcgtgagccaccccgctccaacttcaactcgttgaacttcaagatccccaagcg 2703  
 Qy 1587 ccaactgcccgtg---agtcgaagccttalgaggtcgtgctcgaagagatgtgtcgtgc 1643  
 Db 2704 ctggtgtgtcgttgccaggtcatgacctagagtgctgtctgtcttaagagatgtcgtctg 2763  
 Qy 1644 tggctcgtcgttgaggtgtgtatgttgagggccaactagccttcccgtaagtctcg 1703  
 Db 2764 tgggcaaggggtgagagcccgagacatcgagggccaactaaccacccgtgagc----- 2812  
 Qy 1704 ccttlaactaactgttttcaactatgatacatatcaagaatgtgttcatatgttga 1763  
 Db 2813 -----tgagacttataatgtgca 2830  
 Qy 1764 ctgcacaacactgtatccagagagcaacagatgagctgtcttcaatgtcaatgttct 1823  
 Db 2831 ctgtcaacaacctcatcagagatlaacgaactgagctgatatcaacgaccccat 2890  
 Qy 1824 cgtgtactatgtctacaactacaacgaggttaattgaccccatggagcctctctgagagcc 1883

Db 2891 ggaagagaagagatatcttcaaggaagacttcgagagaccccaatgaacccaagtgtgcgc 2950  
 Qy 1884 ccgccccttccctcctcctcgcgagaggttcgagaatgtctcgtggtgacttcaagcgttgc 1943  
 Db 2951 cgttcccttaaacgcgaacgacttccatgtctcgtggaacttccctcgcgaggtccat 3010  
 Qy 1944 caactgacgcatcagagagatgtgacttaacccctacgcccaggtcgatga 1997  
 Db 3011 cactgcccaggtgcagaggtgtgcgagcagagagccgtacacccgctcgtatga 3064

Search completed: October 5, 2002, 20:29:54  
 Job time: 9023 sec





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 18:54:14 ; Search time 3211.21 Seconds

(without alignments)  
8805.441 Million cell updates/sec

Title: US-09-656-640A-3

Perfect score: 2095  
Sequence: 1 cagctcggtctactctc.....gaataagttgtgtgcttaa 2095

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estbta:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estcov:\*  
6: em\_estipl:\*  
7: em\_estiro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.4	4.1	600	BE188240	M7ATG In
2	86.4	4.1	700	BE187716	BI1809 In
3	68.6	3.3	664	BE188099	CRC377-R
4	65.6	3.1	848	BE216983	EST0376 T
5	63.4	3.0	880	BE264805	HF_CEA001
6	54.6	2.6	465	BC053739	BC053739 RH12.9_B
7	51.4	2.5	311	AV426401	AV426401
8	50	2.4	580	BE159802	BE159802
9	48.2	2.3	592	BE143324	BE143324
10	47.8	2.3	766	BC526127	BC526127 56-88 Ste
11	46.2	2.2	286	A2714471	A2714471 RPCR-24-1
12	46.2	2.2	599	BE258534	BE258534 HVSMEF001
13	46.2	2.2	669	BE258534	BE258534 HVSMEF001
14	45.6	2.2	485	BE1876848	BE1876848
15	45.6	2.2	562	BE651243	BE651243
16	45.6	2.2	596	BE1574420	BE1574420 uk20h11.Y
17	45.6	2.2	600	BE634973	BE634973

18	45.6	2.2	634	BE1574382	BE1574382
19	45.6	2.2	738	BE918653	BE918653
20	45.6	2.2	762	BE1516242	BE1516242
21	44.8	2.1	657	BE1527303	BE1527303
22	44.2	2.1	705	BE1590982	BE1590982
23	43.8	2.1	433	BE018870	BE018870
24	43	2.1	552	BE2302933	BE2302933
25	42.8	2.0	473	BE2302933	BE2302933
26	42.8	2.0	542	BE2302933	BE2302933
27	42.8	2.0	611	BE2302933	BE2302933
28	42.6	2.0	725	BE2302933	BE2302933
29	42.4	2.0	533	BE2302933	BE2302933
30	42.4	2.0	551	BE2302933	BE2302933
31	42.4	2.0	591	BE2302933	BE2302933
32	42.4	2.0	697	BE2302933	BE2302933
33	42.4	2.0	795	BE2302933	BE2302933
34	42.4	2.0	885	BE2302933	BE2302933
35	41.8	2.0	667	BE212100	BE212100
36	41.6	2.0	444	BE100862	BE100862
37	41.6	2.0	518	BE100862	BE100862
38	41.2	2.0	339	BE255913	BE255913
39	41.2	2.0	541	BE255913	BE255913
40	41.2	2.0	722	BE255913	BE255913
41	41.2	2.0	828	BE255913	BE255913
42	41.2	2.0	1128	BE255913	BE255913
43	40.8	1.9	518	BE031325	BE031325
44	40.8	1.9	534	BE031325	BE031325
45	40.6	1.9	316	BE725648	BE725648

#### ALIGNMENTS

RESULT 1  
LOCUS BE188240 600 bp mRNA linear EST 25-SEP-2000  
DEFINITION M7ATG In vitro expressed cDNAs Cladosporium fulvum similar to  
Bilirubin oxidase precursor. mRNA sequence.  
ACCESSION BE188240  
VERSION BE188240.1 GI:8667479  
KEYWORDS  
SOURCE Cladosporium fulvum.  
ORGANISM Cladosporium fulvum  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et  
Chaetothyrionmycetes Incertae sedis; Mycosphaerellaceae; mitosporic  
Mycosphaerellaceae; Cladosporium.  
1 (bases 1 to 600)  
REFERENCE  
AUTHORS Clark,A.J., Rasmussen,S.W. and Oliver,R.P.  
TITLE In vitro expressed genes of Cladosporium fulvum  
JOURNAL Unpublished (2000)  
COMMENT  
Contact: R.P.Oliver  
Neurotrophic Phytopathology Research Centre  
Murdoch University  
SABC, Perth 6150, Western Australia  
Tel: +61-8-9360-7404  
Fax: +61-8-9360-6303  
Email: roliver@central.murdoch.edu.au  
High quality sequence stop: 600  
POLYA-NO.

#### FEATURES

##### source

1..600  
/organism="Cladosporium fulvum"  
/strain="Race 4"  
/db\_xref="taxon:5499"  
/clone\_lib="In vitro expressed cDNAs"  
/tissue\_type="Mycelium"  
/note="Vector: lambda bluescript; A mixture of mycelial  
cultures grown in liquid B5 for 48 hours and transferred  
for 24 hours to media lacking carbon, nitrogen or  
supplemented with hydrogen peroxide"  
BASE COUNT 131 a 176 c 162 g 128 t 3 others  
ORIGIN

Query Match 4.1%; Score 86.4; DB 9; Length 600;  
Best Local Similarity 55.0%; Pred. No. 4.8e-11;  
Matches 243; Conservative 0; Mismatches 178; Indels 21; Gaps 3;

```
OY 1253 gagttgataaacactgacgaaggtcattcgatctgctgttgatgaagaatccttgatgcgcc 1312
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 GAGTGCATTAACACCGACACAGGTGATGAGTGTGCTGGAGA-----C 108
OY 1313 gacactctgaggtgctgctcccaacctcgaagatgttccttccccgaaggcggaactgg 1372
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 AGCGTACCGACGAGATGCCAACACAGTACCATCACTCTGAACGGTGCATCGACTGG 168
OY 1373 gaccccgaaaccccaactgatagaagacttcaactc---ggccgtgctaatagacag 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 CCAGCGCAGACGACACTATTGACAAGACCTTCAACTTCAGATGGCGCGGCGGACGTG 228
OY 1430 tggacaataacagaggttcccttcctgcatgctgcgaagaacggtctgctcgcaatgtgcgc 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 TGGACCGTCAATGGCGTGGATTTCAGCGACCCGAACTCGCTGTGCTGAGGCGCA 288
OY 1490 cgcgaacactgtgagatcctgagacttgagaaacactcaacggttgagactcaacctgtt 1549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 CAAGGACACTGTGAGAGGTGGCGCTCTGTCGACACTGGGCGTGTCCAGCCAGTA 348
OY 1550 caactacactcgttgacttcgagtccttccttcgttc---cactgcccgtgagtcgag 1606
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 CATATCCATCTGTGTCAACATGACAGTCTCTGCGCTGCTGCTGCGGCTCTCATG 408
OY 1607 cctttagagctgctgctgctcaagatgtgtctgctgctgctgctgctgagtgctat 1666
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 409 CCATACGAAGCTGCTGCTGTGAAGAAGCTGTATCTGCTGCTGCGAGATTGTGAC 468
OY 1667 gttgagggccactacgctcctt 1688
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 469 GTGCTTGCCCTTTACGCTCTT 490
```

RESULT 2  
BE187716 700 bp mRNA linear EST 25-SEP-2000  
LOCUS BE187716  
DEFINITION BILIOX In vitro expressed cDNAs Cladosporium fulvum cDNA similar to  
Bilirubin oxidase precursor, mRNA sequence.  
ACCESSION BE187716  
VERSION BE187716.1 GI:8666955  
KEYWORDS EST.  
SOURCE Cladosporium fulvum  
ORGANISM Cladosporium fulvum  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et  
Chaetothyriomycetes incertae sedis; Mycosphaerellaceae; mitosporic  
Mycosphaerellaceae; Cladosporium.  
REFERENCE 1 (bases 1 to 700)  
AUTHORS Clark,A.J., Rasmussen,S.W. and Oliver,R.P.  
TITLE In vitro expressed genes of Cladosporium fulvum  
JOURNAL Unpublished (2000)  
COMMENT Contact: R.P.Oliver  
Neurotrophic Phytopathology Research Centre  
Murdoch University  
SABC, Perth 6150, Western Australia  
Tel: +61-8-9360-7404  
Fax: +61-8-9360-6303  
Email: roliver@central.murdoch.edu.au  
High quality sequence stop: 700  
POLYA-No.

FEATURES  
SOURCE Location/Qualifiers  
1. 700  
/organism="Cladosporium fulvum"  
/strain="Race 4"  
/db\_xref="taxon:5499"  
/clone\_lib="In vitro expressed cDNAs"  
/tissue\_type="Mycelium"  
/note="Vector: Lambda bluescript: A mixture of mycelial  
cultures grown in liquid B5 for 48 hours and transferred";

for 24 hours to media lacking carbon, nitrogen or  
supplemented with hydrogen peroxide 7 others  
BASE COUNT 158 a 206 c 183 g 146 t

Query Match 4.1%; Score 86.4; DB 9; Length 700;  
Best Local Similarity 55.0%; Pred. No. 5.3e-11;  
Matches 243; Conservative 0; Mismatches 178; Indels 21; Gaps 3;

```
OY 1253 gagttgataaacactgacgaaggtcattcgatctgctgttgatgaagaatccttgatgcgcc 1312
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 GAGTGCATTAACACCGACACAGGTGATGAGTGTGCTGGAGA-----C 108
OY 1313 gacactctgaggtgctgctcccaacctcgaagatgttccttccccgaaggcggaactgg 1372
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 AGCGTACCGACGAGATGCCAACACAGTACCATCACTCTGAACGGTGCATCGACTGG 168
OY 1373 gaccccgaaaccccaactgatagaagacttcaactc---ggccgtgctaatagacag 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 CCAGCGCAGACGACACTATTGACAAGACCTTCAACTTCAGATGGCGCGGCGGACGTG 228
OY 1430 tggacaataacagaggttcccttcctgcatgctgcgaagaacggtctgctcgcaatgtgcgc 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 TGGACCGTCAATGGCGTGGATTTCAGCGACCCGAACTCGCTGTGCTGAGGCGCA 288
OY 1490 cgcgaacactgtgagatcctgagacttgagaaacactcaacggttgagactcaacctgtt 1549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 CAAGGACACTGTGAGAGGTGGCGCTCTGTCGACACTGGGCGTGTCCAGCCAGTA 348
OY 1550 caactacactcgttgacttcgagtccttccttcgttc---cactgcccgtgagtcgag 1606
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 CATATCCATCTGTGTCAACATGACAGTCTCTGCGCTGCTGCTGCGGCTCTCATG 408
OY 1607 cctttagagctgctgctgctcaagatgtgtctgctgctgctgctgctgagtgctat 1666
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 409 CCATACGAAGCTGCTGCTGTGAAGAAGCTGTATCTGCTGCTGCGAGATTGTGAC 468
OY 1667 gttgagggccactacgctcctt 1688
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 469 GTGCTTGCCCTTTACGCTCTT 490
```

RESULT 3  
BE188099 664 bp mRNA linear EST 25-SEP-2000  
LOCUS BE188099  
DEFINITION CFC377-R In vitro expressed cDNAs Cladosporium fulvum cDNA similar  
to Bilirubin oxidase precursor, mRNA sequence.  
ACCESSION BE188099  
VERSION BE188099.1 GI:8667338  
KEYWORDS EST.  
SOURCE Cladosporium fulvum  
ORGANISM Cladosporium fulvum  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et  
Chaetothyriomycetes incertae sedis; Mycosphaerellaceae; mitosporic  
Mycosphaerellaceae; Cladosporium.  
REFERENCE 1 (bases 1 to 664)  
AUTHORS Clark,A.J., Rasmussen,S.W. and Oliver,R.P.  
TITLE In vitro expressed genes of Cladosporium fulvum  
JOURNAL Unpublished (2000)  
COMMENT Contact: R.P.Oliver  
Neurotrophic Phytopathology Research Centre  
Murdoch University  
SABC, Perth 6150, Western Australia  
Tel: +61-8-9360-7404  
Fax: +61-8-9360-6303  
Email: roliver@central.murdoch.edu.au  
High quality sequence stop: 664  
POLYA-No.

FEATURES  
SOURCE Location/Qualifiers  
1. 664  
/organism="Cladosporium fulvum"  
/strain="Race 4"

```

/db.xref="taxon:5499"
/clone.lib="In vitro expressed cDNAs"
/tissue_type="Mycellium"
/note="Vector: Lambda bluescript; A mixture of mycelial
cultures grown in liquid B5 for 48 hours and transferred
for 24 hours to media lacking carbon, nitrogen or
supplemented with hydrogen peroxide"
BASE COUNT      149 a      174 c      174 g      140 t      27 others
ORIGIN

Query Match      3.3%; Score 68.6; Db 9; Length 664;
Best Local Similarity 58.1%; Pred. NO.1.6e-06;
Matches 158; Conservative 0; Mismatches 107; Indels 7; Gaps 2;

QY 952 gaacgctgaacgcttgatgctcaaaagctgacgcgcgaagtaacgcttcgctcct 1011
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 57 GAACGCGAGCGCAGCGCATATCTCAAGCTCGAGCGCTGCGAAGTACCATTCAGACTCTT 116

QY 1012 caacgctgcgcltcacgctcttcgctctgatatctgctacacctgagatltcagagac 1071
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 117 TCATATCTCTCTTAAGCAAGATCTTACGACATCTACATTCG-----CTGATGAGAACACAA 170

QY 1072 cagacttcccttcagagctatccgcgcgagagtgctgtcttgaggacctgttgaac 1131
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 171 CCCATCCCAATTCACGAGCTATGCTCAGACTCCGCGCTATTCCGCCGCTNCCGTGAGCGC 230

QY 1132 tgacactctgcatcctctatgcccgaagcgcgtggaggtgttatcagactctccacct 1191
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 231 CAAGACCTGGTGCATTTCCATTCAGGAGACGCT-TCAGATCATTTATTCAGATTATCTGTTT 289

QY 1192 cgcctggccagctcatcgatatccgcgaaccttc 1223
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 290 CGCGGCGCAGATATTACTATGTTTCANCGGTC 321

RESULT 4
LOCUS      BE216983      848 bp      mRNA      linear      EST 03-JUL-2000
DEFINITION EST0376 Triticum aestivum Lambda zap Triticum aestivum cDNA clone
JAI_5A_A06_T3 5', mRNA sequence.
BE216983
ACCESSION BE216983
VERSION   BE216983.1 GI:8904523
KEYWORDS  EST.
SOURCE    bread wheat.
          Triticum aestivum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
          ; Triliceae; Triticum.
          1 (bases 1 to 848)
REFERENCE Anderson,J.M., Williams,C.E. and Goodwin,S.B.
AUTHORS   Analysis of an EST database reveals a probable CR2 resistance gene
TITLE     homolog in wheat
          Unpublished (2000)
JOURNAL   COMMENT
          Contact: Anderson, J.M.
          Crop Production & Pest Control Research Unit
          USDA-ARS
          1150 Lilly Hall, West Lafayette, IN 47907, USA
          Tel: 765-494-5565
          Fax: 765-496-2926
          Email: janderson@purdue.edu
          Seq primer: T3
          High quality sequence stop: 848.
          Location/Qualifiers
            1..848
              /organism="Triticum aestivum"
              /strain="P29"
              /db_xref="taxon:4565"
              /clone="JAI_5A_A06_T3"
              /clone.lib="Triticum aestivum Lambda zap"
              /tissue_type="leaf"
              /dev_stage="9 day old seedlings"
              /dev_stage="9 day old seedlings"
              312 c      263 g      134 t
BASE COUNT      139 a      263 g      134 t

```

	ORIGIN	
Query Match	3.1%; Score 65.6; DB 9; Length 848;	
Best Local Similarity	55.2%; Pred. No. 1,1e-05;	
Matches 128; Conservative	0; Mismatches 104; Indels 0; Gaps 0;	
OY	550 gcttcggcagatcaagaagttaacttaccaccaacaggcaggtctgccgcgatglttgta 649	
Dd	1 GCACAGGACTTCGCCGGCGGTACGAGTACCACAGCAGACGCTCCGGTAACTCTTGSTA 60	
OY	650 ccatgacctatgcattctcatccacgcgcgcgcgcgaagccctaatggttgtagtgttcta 709	
Dd	61 CCACAGCACACGCATGGGAGCTCATCCCGGCTCAACATCTCTGCTGGCTTCCTGGGGCGTA 120	
OY	710 catgtccaaagaccgcgcgtctgtagatgacctgaacctcccacgcgcgcgcgcgcgtttga 769	
Dd	121 CCGTCTCCCCCAGCACC GGCGAGCCGAGGCCGCCCTCGGCTCCCTCCGGCGGGAATTGCA 180	
OY	770 tatccccctgttctctgactgccaagcagatacaaacgacagcgactctctc 821	
Dd	181 CCGAACACTCGHCTCTCTTCAGACCGGCACTTCAAGAGCAGCGGCGGCTCTTC 232	
RESULT 5		
LOCUS	BF264805	880 bp mRNA linear EST 23-OCT-2001
DEFINITION	HV.CEA0010G19f Hordeum vulgare seedling green leaf EST library HVCDNA0004 (Blumeria challenged) Hordeum vulgare CDNA clone	
ACCESSION	BF264805	
VERSION	BF264805.2 GI:13261741	
KEYWORDS	EST.	
SOURCE	Barley.	
ORGANISM	Hordeum vulgare	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Hordeum. 1 (bases 1 to 880)	
AUTHORS	Wang,R., Close,T.J., Kleinhofs,A., Wise,R., Wei,F., Begum,D., Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi, D.W., Fenton,R.D., Oates,R., and Main,D. Development of a genetically and physically anchored EST resource for barley genomics: Blumeria infected incompatible (Mla13) seedling leaf CDNA library Unpublished (2001)	
JOURNAL COMMENT	On Nov 17, 2000 this sequence version replaced gi:11195799.  Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Total hg bases = 260 Seq primer: AATTACCTCTCACTTAAGG High quality sequence stop: 798.	
FEATURES	Location/Qualifiers	
SOURCE	1..880	
	/organism="Hordeum vulgare"	
	/cultivar="CI16155 (Mla13)"	
	/db_xref="taxon:4513"	
	/clone="HV.CEA0010G19f"	
	/clone.lib="Hordeum vulgare seedling green leaf EST library HVCDNA0004 (Blumeria challenged)"	
	/russus_type="Seedling green leaf"	
	/lab_host="TJC121"	
	/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; C.I. 16155 (Mla13) plants were greenhouse grown in the R wise lab at Iowa State University, Ames, IA; 7 day old green seedlings were challenged with isolate A27 (AVR-Mla13 ) of Blumeria graminis f. sp. hordei, and leaves were harvested 20 and 24 hr post-inoculation and snap frozen;	

uninoculated leaves were harvested 20 hr post-inoculation (Wei, Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool. Equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one cDNA library was made, and 1 million pfu were in vivo excluded to give pluescript SK(-) cDNA phagemids (Choi, Close). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmett, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Math). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see

<http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinof's A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 215 a 198 c 318 g 149 t

Query Match 3.0% Score 63.4; DB 10; Length 880;

Best Local Similarity 55.8%; Pred No. 4.1e-05;

Matches 121; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

OY 605 gattactactaccacaaagagcgtgcccagatcttggttaccatgacatgcat 664

DB 74 ggctgacgagatgacccgacagcagcagcgtgacgtgacgacgacgacgac 133

OY 665 gctacacacgcgcgagaaagcgtacatggtgacgtggtctacatgacgagcc 724

DB 134 ggccctcaccgctgacacacacacacacacacacacacacacacacacac 193

OY 725 gactgagagatgctgacacacacacacacacacacacacacacacacacac 784

DB 194 ggccacgagagggggccctgctgctgctgctgctgctgctgctgctgctgct 253

OY 785 gactgacacgac 821

DB 254 cttgcacgacgac 290

RESULT 6

LOCUS BG053739 465 bp mRNA linear EST 25-JAN-2001

DEFINITION RH12.9\_B08.D1\_A003 Rhizome2 (RH12) Sorghum propinquum cDNA, mRNA

ACCESSION BG053739

VERSION BG053739.1 GI:12509741

KEYWORDS EST

SOURCE Sorghum propinquum

ORGANISM Sorghum propinquum

REFERENCE 1 (bases 1 to 465)

AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Paterson, A., Sudman, M. and Pratt, L.H.

TITLE An EST database from Sorghum: Sorghum propinquum rhizomes

JOURNAL Unpublished (2000)

COMMENT Department of Botany

The University of Georgia

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 542 1805

Email: mmpratt@uga.edu

High quality sequence stop: 421

POLYA-No.

FEATURES

source

Location/Qualifiers

1..465

/organism="Sorghum propinquum"

/db\_xref="taxon:132711"

/clone\_lib="Rhizome2 (RH122)"

/note="Organ: Rhizomes; Vector: pluescript II from lambda

zap II; Site:1: XhoI; Site:2: EcoRI; The library was made

from poly-A RNA in the cloning vector lambda zap II."

clones to be sequenced were prepared by mass excision."

BASE COUNT 72 a 128 c 193 g 72 t

ORIGIN

Query Match 2.6% Score 54.6; DB 10; Length 465;

Best Local Similarity 56.8%; Pred. No. 0.0046;

Matches 121; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

OY 609 tactactaccacaaagagcgtgcccagatcttggttaccatgacatgcatgcat 668

DB 452 TACCGGTACCCACACGTCGACCGCGGACACCTGTGTACACACACACGCGCTGCGC 393

OY 669 atcaccgacgagacgctacatggtgacgtggtgtacatgacacgagaccgagct 728

DB 392 CTCACCCGCGCCACACCTCTCGCGGCGCTCTGCGGCGCTACGTCATGAGAAAGCGGAG 333

OY 729 gagagatgcttgaacacacacacacacacacacacacacacacacacac 788

DB 332 GTCGACGTCCTCATGATCTTCTGCGACGACGACGACGACGACGACGACGAC 276

OY 789 gccaaagacatcacagcagcagcagcagcagcagcagcagcagcagcagc 821

DB 275 GACCGACGCTTACGTCGACGCGCTGCTGTAC 243

RESULT 7

LOCUS AV426401 311 bp mRNA linear EST 23-MAY-2000

DEFINITION AV426401 Lotus japonicus young plants (two-week old) Lotus

japonicus cDNA clone MM06f04\_f\_5', mRNA sequence.

ACCESSION AV426401

VERSION AV426401.1 GI:7785301

KEYWORDS EST

SOURCE Lotus japonicus

ORGANISM Lotus japonicus

REFERENCE 1 (bases 1 to 311)

AUTHORS Asanizu, E., Nakamura, Y., Sato, S. and Tabata, S.

TITLE Generation of 7137 non-redundant expressed sequence tags from a

legume, Lotus japonicus

JOURNAL DNA Res. 7 (2), 127-130 (2000)

COMMENT Contact: Yasukazu Nakamura

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yama 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: ynakamu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

Location/Qualifiers

1..311

/organism="Lotus japonicus"

/db\_xref="taxon:34305"

/clone\_lib="MM06f04\_f\_5"

/dev\_stage="young plants (two-week old)"

/note="Vector: pluescript II SK-; Site:1: EcoRI; Site:2: XhoI; isolate-Miyakojima MG-20"

clones to be sequenced were prepared by mass excision."

BASE COUNT 81 a 100 c 68 g 62 t

ORIGIN



BASE COUNT  
ORIGIN

155 a 177 c 136 g 124 t

/dev\_stage="2-3 weeks old"  
/lab\_host="DH10B"  
/note="Vector: p Bluescript II SK+, Site\_1: EcoRI, Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from 2-3 week old seedlings that were induced for symptoms of SDS (Sudden Death Syndrome) disease by the translocation of culture filtrate of *Fusarium solani* f. sp. *glycines* (Plant Cell Report 18:375-380). Cultivar PI 567374 is partially resistant to the disease SDS. Plant tissue (expanded leaves, folded leaves, and new shoots) were collected at 1, 6, 24, and 48 hrs. after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Strategene p Bluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the p Bluescript vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Steve Clough (Lila Vodka lab, University of Illinois)."

Query Match	2.3%	Score 48.2;	DB 10;	Length 592;
Best Local Similarity	51.1%;	Pred. No. 0.22;		
Matches 113; Conservative	0;	Mismatches 108;	Indels 0;	Gaps 0;

Accession	Sequence	Position
Oy	acaagatctactatcaccccaagcagcagctgcccgaatcttggtacacatbaccat	60
Oy		61
Oy		62
Oy		63
Oy		64
Oy		65
Oy		66
Oy		67
Oy		68
Oy		69
Oy		70
Oy		71
Oy		72
Oy		73
Oy		74
Oy		75
Oy		76
Oy		77
Oy		78
Oy		79
Oy		80
Oy		81
Oy		82
Oy		83
Oy		84
Oy		85
Oy		86
Oy		87
Oy		88
Oy		89
Oy		90
Oy		91
Oy		92
Oy		93
Oy		94
Oy		95
Oy		96
Oy		97
Oy		98
Oy		99
Oy		100
Oy		101
Oy		102
Oy		103
Oy		104
Oy		105
Oy		106
Oy		107
Oy		108
Oy		109
Oy		110
Oy		111
Oy		112
Oy		113
Oy		114
Oy		115
Oy		116
Oy		117
Oy		118
Oy		119
Oy		120
Oy		121
Oy		122
Oy		123
Oy		124
Oy		125
Oy		126
Oy		127
Oy		128
Oy		129
Oy		130
Oy		131
Oy		132
Oy		133
Oy		134
Oy		135
Oy		136
Oy		137
Oy		138
Oy		139
Oy		140
Oy		141
Oy		142
Oy		143
Oy		144
Oy		145
Oy		146
Oy		147
Oy		148
Oy		149
Oy		150
Oy		151
Oy		152
Oy		153
Oy		154
Oy		155
Oy		156
Oy		157
Oy		158
Oy		159
Oy		160
Oy		161
Oy		162
Oy		163
Oy		164
Oy		165
Oy		166
Oy		167
Oy		168
Oy		169
Oy		170
Oy		171
Oy		172
Oy		173
Oy		174
Oy		175
Oy		176
Oy		177
Oy		178
Oy		179
Oy		180
Oy		181
Oy		182
Oy		183
Oy		184
Oy		185
Oy		186
Oy		187
Oy		188
Oy		189
Oy		190

RESULT	10
LOCUS	BG526127
DEFINITION	BG526127 766 bp mRNA linear EST 16-NOV-2001 36-88 <i>Stevia</i> field grown leaf cDNA <i>Stevia rebaudiana</i> CDNA 5', mRNA sequence.

ACCESSION	BG526127
VERSION	BG526127.1
KEYWORDS	EST.
SOURCE	<i>Stevia rebaudiana</i> .
ORGANISM	<i>Stevia rebaudiana</i>

REFERENCE  
I (pages I to 766)

**AUTHORS** Brande, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.  
**TITLE** Leaf Esters from *Stevia rebaudiana*: A resource for gene discovery in diterpene synthesis  
**JOURNAL** unpublished (2001)  
**COMMENT** Contact: Jim Brande

Genomics and Biotechnology  
Agriculture and Agri-Food Canada - SCPFR

1391 Sandford St., London, Ontario, Canada, N5V 4T3  
Tel: 519 457 1470  
Fax: 519 457 3897  
Email: brandtje@agr.ca  
Seq primer: T3 promoter primer.  
Location/Qualifiers  
1. 766

BASE COUNT	203 a	192 c	174 g	191 t	6 others
ORIGIN					

	Query Match	2.3%	Score 47.8	DB 10	Length 766					
	Best Local Similarity	52.9%	Pred. No. 0.33							
	Matches 100	Conservative	0	Mismatches 89	Indels 0	Gaps 0				
QY	601	acaagat	tactactacc	cccaacagcagc	gctgcgcga	tgcctt	gtatccatg	accatg	660	
Db	64	AAANNCCT	TACACATCA	CAACACAA	CAACCCGCG	CAACCTTAT	TACCTGATG	ATCAAG	123	
QY	661	ccatctccat	caaccgcg	agaaagcc	tcaatcgtgc	tcagctgt	gtctatca	atgatcc	720	
Db	124	CCATGGGGT	TGACCA	GGGGCAAC	CTTCTT	TGACAG	TTTGTTCG	AGACTTAT	CAATCA	183
QY	721	accgcgc	tgcagatgc	cccttgaa	ccctccca	gcgagct	tcagcgc	agtttga	tatccct	780
Db	184	ATCACAG	TATGCAAG	CCCACTT	GGGTC	CTTTCG	TGATGAT	GATTTG	ATGCGCG	243
QY	781	tttcgactg	789							
Db	244	TCGTATTG	252							

RESULT	11	1
LOCUS	AZ714471	
DEFINITION	RPCT-24-138E2.TJ RPCT-24 Mus musculus genomic clone RPCT-24-138E2, GSS 24-JAN-2001	

ACCESSION	A2714471
VERSION	A2714471.1
KEYWORDS	GSS.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE  
AUTHORS  
Zhao, S., Niernan, W., Malek, J., Shatsman, S., Aktinety, B., Levins, M.,

Tsegaye, G., Geer, K., Krol, M., Shwartsbeyn, A.  
Russell, D., de Jong, P. and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-20  
Unpublished (1999)

TITLE Mouse BAC End Sequences from Library RPCI-24  
JOURNAL Unpublished (1999)







```

/sex="female"
/dev stage="adult"
/lab host="DH10B"
/notes="Organ: kidney. Vector: pME18S-FL3. Site_1: DraIII
(CACGTGTGTG), Site_2: DraIII (CACCATGTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGAGGACCTTCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGTGACCTGTGTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACGTGTGTG, 3' site CACCATGTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTTAAAGAGCTCG and 3' end
primer GGACCTCCACATCTCAGACAC."
122 a 137 c 120 g 106 t

```

Query Match	2.2%	Score 45.6;	DB 9;	Length 485;
Best Local Similarity	52.0%;	Pred. No. 0.9;		
Matches 102; Conservative	0;	Mismatches 94;	Indels 0;	Gaps 0

**QY** 542 ctccttctcgcagcccccctttgatgttgctgtagaacactaccagcctgycgagt 601  
|| ||| ||| ||| ||| ||| ||| |||  
**Db** 265 CTCCTGTGCAAGTGGCAGSAGATGCCAAGAATAAGSACATCTCCTGTTCCTCCCCAA 324

Oy 602 caaggattactactaccccaacagcgagcgtgccgcgcatgcttggtaacatgatgc 661  
||| | ||| ||| ||| | ||| ||| |  
Db 325 TGGGGAAGAACTGAGCCCAACACCAGCAGCGGATCTCAGTGGGTGTGGAATGATGACGACTC 384

QY 662 catgtccatcaacgcgcgagaaocctcatgtgtcaggctgtgtctacatgatccaga 721  
| | | | | | | | | | | | | | |  
Db 385 CTCTACCCCTACCATCTTACAAGGCCACACATCGACGATGCCGGCATATACAAGTGTTGTG 444

QY 722 ccgcgtgagatgcc 737  
| | | | | | | | | |  
Db 445 CACGGCTGAGGACGGC 460

RESULT	15
LOCUS	BB651243
DEFINITION	BB651243 RIKEN full-length enriched, 0 day neonate cerebellum mus musculus cDNA C230070H21 5', mRNA sequence.

VERSION	BB651243.1	GI:16485497
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE  
AUTHORS  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

TITLE  
JOURNAL  
COMMENT

Yoshihide Hayashizaki

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Mimoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, H., Koyama, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, H., Okazaki, Y., Okido, T., Saito, R., Sakai, K., Sano, H., Sasakawa, D., Shibata, K., Shindagawa, A., Shirai, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M., and Hayashizaki, T.

RIKEN MOUSE ESIS (Arakawa, T., et al. 2001)

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Saitiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel.: 81-45-503-9322  
Fax: 81-45-503-9316  
Email: genome.res@gsc.riken.go.jp/  
URL: <http://genome.gsc.riken.go.jp/>  
carlinci, P., Shibata, Y., Hayashi, N., Sugahara, Y., Shibata, K., Itoh

'M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E.,  
 Matabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
 'S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
 10 (11), 1757-1771 (2000)  
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 'Y., and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
 Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa  
 'K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences. *Mamm. Genome.* 12, 673-677 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.  
 e mouse tissues.  
 Location/Qualifiers  
 1. 562

FEATURES

SOURCE

```

/db_xref="taxon:10090"
/clone="C230070H21"
/clone_lib="RIKEN full-length enriched, 0 day neonate

```

```
/tissue_type="cerebellum"  
/dev_stage="0 day neonate"  
/lab_host="DH103"
```

prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in

...cagcaaaacatggttttctgggaacac  
contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']  
GAGGAGACAGCATCCAAAGACTCTTTTTTTTTTTTNN 3'|, cDNA was

Preparation of coding chromosomal DNA extracted, religated and subsequently enriched for full-length transcripts. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 479.0. Second

sequence [5' GAGACAGACATTTCTCGATTAAATTAATTAATCCCCCCCCCCCCCCCCCC  
3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified plasmidscript KS(+) after bulk excision from lambda<sub>gt</sub>10.

BASE COUNT	141 a	160 c	142 g	119 t
ORIGIN				

Query Match	2.2%	Score	45.6	DB	9	length	562
Best Local Similarity	52.0%	Pred. No.	0.99				
Matches 102; Conservative	0	Mismatches	94	Indels	0	Gaps	0

Qy 542 cctcttctctcgagctcccttttgatggttgctgtagaacactaccacgcctgcgagta 601  
||| ||| | |||| | ||||| ||| |  
Db 303 CTTCCTGTGTCAGATGGCAGAGATGCCAAGAATTAGGACATCTCCTTGTTCTCCCCCA 362

Oy 602 caagatctactactaccaccaagcgcgctgcgcgcatgtcttgytaccatgacatgc 661  
 ||| | ||| ||| ||| | ||| ||| | ||| | |  
 Db 363 TGGGGAGAAAGCTGAGCCCAACACGACGCGATCTCAGTGGGTGGGAATGATGACACACTC 422

QY 662 catgtccatcacccgcgcagaaagccctacatggtgcagctgtgtctacatgatccagga 721  
 1 11 11111 1 111111 1111 1 111 1 1111 1 11  
 Db 423 CTCTACCCCTCACCACCTTACCAACGCCACCAATGCATGCATGCGGCATATACAAAGTGGTGGC 482

722 ccgcgtgagatgcc 737  
| | | | | | | | | |  
QY

Mon Oct 7 11:28:44 2002

us-09-656-640a-3.rst

Page 10

Db 483 CACGGCTGAGAGGCC 498

Search completed: October 5, 2002, 18:54:25  
Job time: 3515 sec

---



OY	421	tgrtctgtatcagatgagcatgtccccagagactcaactcaatcgtttccctcgttgcacttgaga	480
Db	421	tgrtctgtatcagatgagcatgtccccagagactcaactcaactcgtttccctcgttgcacttgaga	480
OY	481	gtgtgtgtccgcttcgttgaaacgcggaagagaacacctctccaaacgcgtgtccacttgcagc	540
Db	481	gtgtgtgtccgcttcgttgaaacgcggaagagaacacctctccaaacgcgtgtccacttgcagc	540
OY	541	gctcttctctcgagactccctctgtatggttggcctgagagacaactaacccagccctggcagat	600
Db	541	gctcttctctcgagactccctctgtatggttggcctgagagacaactaacccagccctggcagat	600
OY	601	acaagagattactatcaccccaacagacgcctgcgcgcacatgttttggatgcatagcacgt	660
Db	601	acaagagattactatcaccccaacagacgcctgcgcgcacatgttttggatgcatagcacgt	660
OY	661	ccatgtccatcacccgcgcgagagacgcctacatgtggttcagacgttggttctactatgatccag	720
Db	661	ccatgtccatcacccgcgcgagagacgcctacatgtggttctactatgatccag	720
OY	721	accgcgctgtagagatgcccctgaacctctcccaagcgctcagcgcgagtttgatatccctctgg	780
Db	721	accgcgctgtagagatgcccctgaacctctcccaagcgctcagcgcgagtttgatatccctctgg	780
OY	781	tctctgactgcacaagatcgtataacacgcagacgcctctctctccacaacaaatgtagaggttt	840
Db	781	tctctgactgcacaagatcgtataacacgcagacgcctctctctccacaacaaatgtagaggttt	840
OY	841	ccagactctctggggtgtagcgttatctcaagctgtgtgaagttcagaccattgtagatgtctcagatc	900
Db	841	ccagactctctggggtgtagcgttatctcaagctgtgtgaagttcagaccattgtagatgtctcagatc	900
OY	901	ctagaagatattcgaatgtgaatgttgcaatgctctacaacagtgctataccaaagagtgca	960
Db	901	ctagaagatattcgaatgtgaatgttgcaatgctctacaacagtgctataccaaagagtgca	960
OY	961	gctcttggtcctatgtctcaacgctgcagccgcgcaagtaacgcgtctccgtctctcaacgctgc	1020
Db	961	gctcttggtcctatgtctcaacgctgcagccgcgcaagtaacgcgtctccgtctctcaacgctgc	1020
OY	1021	cgctccacgcgtctctgcgtctgtatctctgcatacctctgtagagattcagaagacaagacttcc	1080
Db	1021	cgctccacgcgtctctgcgtctgtatctctgcatacctctgtagagattcagaagacaagacttcc	1080
OY	1081	cttccagagctcatctgcgcctgcagcgtgctgcgtcttgagggcctctgttgacaactgacactct	1140
Db	1081	cttccagagctcatctgcgcctgcagcgtgctgcgtcttgagggcctctgttgacaactgacactct	1140
OY	1141	gtatactctctatggccgagcgtcgggaggtttgtatcagactcttccacacttcgcctgcgcga	1200
Db	1141	gtatactctctatggccgagcgtcgggaggtttgtatcagactcttccacacttcgcctgcgcga	1200
OY	1201	gtccatctgatatccgcgaacactctctgcgtctgcagcgtctcccggtgtcttgagccttgagtttga	1260
Db	1201	gtccatctgatatccgcgaacactctctgcgtctgcagcgtctcccggtgtgttgagccttgagtttga	1260
OY	1261	taaacactgcacaaggtctacgcatctgcttgatgtgaatgactcttgagccgcgcgaactctc	1320
Db	1261	taaacactgcacaaggtctacgcatctgcttgatgtgaatgactcttgagccgcgcgaactctc	1320
OY	1321	tgaagtgctgcgcgaacactctcgagatgttctctctcccgaggcgcgacaacttgagcccgcc	1380
Db	1321	tgaagtgctgcgcgaacactctcgagatgttctctctcccgaggcgcgacaacttgagcccgcc	1380
OY	1381	aaaccccaactgatatcagagactcttcaactcttcgcgcgtctgcttaatgtgacaagtgaacaatca	1440
Db	1381	aaaccccaactgatatcagagactcttcaactcttcgcgcgtctgcttaatgtgacaagtgaacaatca	1440
OY	1441	cggaggttaactcttctcgagatgtctgagaacacgctgtctccgcgaatgtgcgcccgcgacaactgt	1500
Db	1441	cggaggttaactcttctcgagatgtctgagaacacgctgtctccgcgaatgtgcgcccgcgacaactgt	1500

QY	1501	tgaagatcggcgacttgaagaacaattccaaacggttggacttaacacctgtttacaattcaact	1500
Db	1501	tgaagatcggcgacttgaagaacaattccaaacggttggacttaacacctgtttacaattcaact	1500
QY	1561	cgttactccgaattcccttcttcgttccaatgccccggagtcgagccctcttgaagctgc	1620
Db	1561	cgttactccgaattcccttcttcgttccaatgccccggagtcgagccctcttgaagctgc	1620
QY	1621	tgtctccaagatgtgtcctgctgctgcgtcgtcgtgaaggttgtctatgtgtgaagcccaata	1660
Db	1621	tgtctccaagatgtgtcctgctgctgcgtcgtcgtgaaggttgtctatgtgtgaagcccaata	1660
QY	1681	cgctctcttcccgaaagtcttcgtcccttttaaccacttggttttaccatgttaacatct	1740
Db	1681	cgctctcttcccgaaagtcttcgtcccttttaaccacttggttttaccatgttaacatct	1740
QY	1741	acaaatgtgtctactaagtttgcgaactgcacaactatcatccaagagacacagacatgtg	1800
Db	1741	acaaatgtgtctactaagtttgcgaactgcacaactatcatccaagagacacagacatgtg	1800
QY	1801	gctgctcttcaatgtcactgttcttcggttgaataatgctacaaactaaacccgagtttcaatgac	1860
Db	1801	gctgctcttcaatgtcactgttcttcggttgaataatgctacaaactaaacccgagtttcaatgac	1860
QY	1861	cccaatggagcgctctcttgagaagccccgcgcctctctctctcttgagaagattcgagaatgctgcg	1920
Db	1861	cccaatggagcgctctcttgagaagccccgcgcctctctctctcttgagaagattcgagaatgctgcg	1920
QY	1921	ggtgactccaagcgagcttgccatcaactagacccgattcagaagatgctgacttcaacccc	1980
Db	1921	ggtgactccaagcgagcttgccatcaactagacccgattcagaagatgctgacttcaacccc	1980
QY	1981	taagcccaagctgatagtatgctgcgtctgagagtaataatgataatgctgctgaatgatatta	2040
Db	1981	taagcccaagctgatagtatgctgcgtctgagagtaataatgataatgctgctgaatgatatta	2040
QY	2041	tggacagcagatataagctatctttgaagaatacttgataaagtgtgtgtgttaa	2095
Db	2041	tggacagcagatataagctatctttgaagaatacttgataaagtgtgtgtgttaa	2095
RESULT 2			
US-09-401-476-1			
; Sequence 1, Application US/09401476			
; Patent No. 6168936			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Huaming			
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes			
; FILE REFERENCE: GCS54			
; CURRENT APPLICATION NUMBER: US/09/401,476			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 1			
; LENGTH: 1958			
; TYPE: DNA			
; ORGANISM: Stachybotrys chararum			
US-09-401-476-1			

Query Match	92.9%	Score 1946.2	DB 4	Length 1958
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1948: Conservative	0	Mismatches 3	Indels 0	Gaps 0
QY 67	aacgatacaacatgatacgaagcaagctatcggacgctgtgctctggagccttgcgtgatacgg	126		
Db 3	atccatcaacaatgatacgaagcaagcatalcggacgctgtgctctggagccttgcgtgatacgg	62		
QY 127	cgagcagctctgcgatacgcagataccgcttgccttgcctgatacgaagacatgaccccttcgagct	186		
Db 63	cgagcagctctgcgatacgcagataccgcttgccttgcctgatacgaagacatgaccccttcgagct	122		
QY 187	caccagaagcagcaagcaagcagctgactctccctctggcctgttgcagaaagtgcctcttcgcat	246		

|||||  
Db 123 caccagaagcagcagcagctgactcccccgccttgtaagagtcccttcgcat 182  
QY 247 cccctctcgaagcgcgcgaagatgaatctatctatagctgaacagcaagcttg 306  
Db 183 cccctctcgaagcgcgcgaagatgaatctatctatagctgaacagcaagcttg 242  
QY 307 ctaatactgcaatgtaacgctcccaacccaacactgagagagacatctgtactacaga 366  
Db 243 ctaatactgcaatgtaacgctcccaacccaacactgagagagacatctgtactacaga 302  
QY 367 tggagatagagccctctcccaacagatcaccctgactgagacgcgcgaacatgttg 426  
Db 303 tggagatagagccctctcccaacagatcaccctgactgagacgcgcgaacatgttg 362  
QY 427 gatacagatgcatgtgcccaagaccatacatctctccctgctgagcactgagatgttg 486  
Db 363 gatacagatgcatgtgcccaagaccatacatctctccctgctgagcactgagatgttg 422  
QY 487 tccgctctgtaacagcgcgcgaagacaacctctcccaacagcgtccactgcaagctct 546  
Db 423 tccgctctgtaacagcgcgcgaagacaacctctcccaacagcgtccactgcaagctct 482  
QY 547 tctctcagctccctctgtaagctgctgagagacatacccaagcctggcgaatagaag 606  
Db 483 tctctcagctccctctgtaagctgctgagagacatacccaagcctggcgaatagaag 542  
QY 607 attactactaccccaacagcgcgcgaagcgcgcgaagcgcgcgcgcgcgcgcgcgcgc 666  
Db 543 attactactaccccaacagcgcgcgaagcgcgcgaagcgcgcgcgcgcgcgcgcgcgc 602  
QY 667 ccataccgcgcgcgaagc 726  
Db 603 ccataccgcgcgcgaagc 662  
QY 727 ctgagagatgcccctgtaacccctcccaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 786  
Db 663 ctgagagatgcccctgtaacccctcccaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 722  
QY 787 ctgagagatgcccctgtaacccctcccaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 846  
Db 723 ctgagagatgcccctgtaacccctcccaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 782  
QY 847 tctgagatgacgttatctcaagctgaagctgaagcccatgagatgtgcttaagctcctaga 906  
Db 783 tctgagatgacgttatctcaagctgaagctgaagcccatgagatgtgcttaagctcctaga 842  
QY 907 gtaacagatgatagaatgtgcaatgctctcaacccagctgatacagagagcgcgcgcgcgc 966  
Db 843 gtaacagatgatagaatgtgcaatgctctcaacccagctgatacagagagcgcgcgcgcgc 902  
QY 967 gtaacagatgatagaatgtgcaatgctctcaacccagctgatacagagagcgcgcgcgcgc 1026  
Db 903 gtaacagatgatagaatgtgcaatgctctcaacccagctgatacagagagcgcgcgcgcgc 962  
QY 1027 acgctcttcgctctgtaacccctgtaacccctgtaacccctgtaacccctgtaacccctgta 1086  
Db 963 acgctcttcgctctgtaacccctgtaacccctgtaacccctgtaacccctgtaacccctgta 1022  
QY 1087 gtaacagatgatagaatgtgcaatgctctcaacccagctgatacagagagcgcgcgcgcgc 1146  
Db 1023 gtaacagatgatagaatgtgcaatgctctcaacccagctgatacagagagcgcgcgcgcgc 1082  
QY 1147 ctctatgagc 1206  
Db 1083 ctctatgagc 1142  
QY 1207 cgaataccgaacactccgc 1266  
Db 1143 cgaataccgaacactccgc 1202  
QY 1267 tgaacagatgatagaatgtgcaatgctctcaacccagctgatacagagagcgcgcgcgcgc 1326  
|||||

Db 1203 tgaacagatgatagaatgtgcaatgctctcaacccagctgatacagagagcgcgcgcgcgc 1262  
QY 1327 gcttgcaacactccgc 1386  
Db 1263 gcttgcaacactccgc 1322  
QY 1387 cactgatacagagactctcaacccctgtaacccctgtaacccctgtaacccctgtaacccctgta 1446  
Db 1323 cactgatacagagactctcaacccctgtaacccctgtaacccctgtaacccctgtaacccctgta 1382  
QY 1447 taactctcgaatgtaacccctgtaacccctgtaacccctgtaacccctgtaacccctgtaac 1506  
Db 1383 taactctcgaatgtaacccctgtaacccctgtaacccctgtaacccctgtaacccctgtaac 1442  
QY 1507 ctgagcacttggaacaaactcccaagcttggaatcaccctgcttaacatcaccctgcttga 1566  
Db 1443 ctgagcacttggaacaaactcccaagcttggaatcaccctgcttaacatcaccctgcttga 1502  
QY 1567 ctccgaatcctctcgc 1626  
Db 1503 ctccgaatcctctcgc 1562  
QY 1627 caagagatgtgtcgc 1686  
Db 1563 caagagatgtgtcgc 1622  
QY 1687 ttcccgtaagctctcgc 1746  
Db 1623 ttcccgtaagctctcgc 1682  
QY 1747 ggtgtacatgcttgcaactgcaacccctgatacagagagacagagacagagacagagacagagac 1806  
Db 1683 ggtgtacatgcttgcaactgcaacccctgatacagagagacagagacagagacagagacagagac 1742  
QY 1807 ttcaatgtcaactgtctcgc 1866  
Db 1743 ttcaatgtcaactgtctcgc 1802  
QY 1867 gtagcctctctgagagc 1926  
Db 1803 gtagcctctctgagagc 1862  
QY 1927 ttcaagcagctctcgaacccctgcaacccctgcaacccctgcaacccctgcaacccctgcaaccc 1986  
Db 1863 ttcaagcagctctcgaacccctgcaacccctgcaacccctgcaacccctgcaacccctgcaaccc 1922  
QY 1987 caagcctgatagatagc 2046  
Db 1923 caagcctgatagatagc 1982  
RESULT 3  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300,6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F15
US-08-232-463-14

```

```

Query Match          2.4% Score 50; DB 1; Length 7218;
Best Local Similarity 0.5%; Pred. No. 0.00012;
Matches 2; Conservative 221; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 1539 ctccacctgtacatctcaccctcgtgacctccgagtcctcttcgttccactgcccgtg 1598
DB 1074 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1133
QY 1599 gactcgagcctatagagctgctgctcgaagatgctcgcgcgcgcgcgcgcgcgcgc 1658
DB 1134 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1193
QY 1659 ttgtctatgttagagccactacgctcctcccgtaagttcctgccttaccactg 1718
DB 1194 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1253
QY 1719 gtttcaactcactgatacactacaaagtgctgctacatgctgacatgccaactgat 1778
DB 1254 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1313
QY 1779 ccacgagaccacgacatgctgcttcaatgctcactgcttctcgtgactatgacta 1838
DB 1314 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1373
QY 1839 caactacacgagatcatgaccacatgagcctcctctgagagcccgcccttccctc 1898
DB 1374 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1433
QY 1899 cggga 1902
DB 1434 ygtta 1437

```

```

RESULT 4
US-07-923-724-1
Sequence 1, Application US/07923724
Patent No. 5780292
GENERAL INFORMATION:
APPLICANT: Nevalainen, Helena K.M.
APPLICANT: Paloneimo, Maria T.
APPLICANT: Miettinen-Oinonen, Arja S.K.
APPLICANT: Torkkeli, Tuula K.
APPLICANT: Cantrell, Michael
APPLICANT: Piddington, Christopher S.
APPLICANT: Ramdasek, John A.

```

```

APPLICANT: Turunen, Maria K.
APPLICANT: Fagerstr m, Richard B.
TITLE OF INVENTION: Production of Phytase Degrading Enzymes
TITLE OF INVENTION: Production of Phytase Degrading Enzymes
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSER: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,724
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/496,155
FILING DATE: 19-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/044,077
FILING DATE: 29-APR-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 8610600
FILING DATE: 30-APR-1986
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 1050,0240004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2071 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: join(136..915, 970..1089, 1142..1245, 1305..1737)
US-07-923-724-1

```

```

Query Match          1.9% Score 39.4; DB 1; Length 2071;
Best Local Similarity 45.6%; Pred. No. 0.088;
Matches 139; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

```

```

QY 455 catcatcgtctcgtgagactgagatgtgtcgccttcgagtcgaagcagcgagagac 514
DB 444 CAGCATCAACTACTGATGATACAGAGGACCTGGCCTTCTGAAGACTGACCTACTA 503
QY 515 ctctccaaacagctccactgacagcgtctctctcgcagctcccttgaatgtgtgagc 574
DB 504 GGTCCCTATATGATGCTACTACTACAGCGCGAGACACAGCGGCGCCCTACGCCGTTTGT 563
QY 575 tgaagacactaccagcctgagctgacaaagattactactaccccaagcagcgtgc 634
DB 564 GAGCCGTTACACACATGGCAACGATTACAGAGCTGCTACGCGCACCTCTGCAACGCTCA 623
QY 635 ccgaatgcttctgtaacatgacatgcatgcatgcatgcatgcatgcatgcatgcatg 694
DB 624 GACGGTGGCCCTCTTTCTAGTGGCTACGAGCTGTCATCGAGAGCGGCCGCAAGTT 683
QY 695 tcaagctgtgtctacatgatccagacccgctgagatgctgacccctgaccccgagcg 754
DB 684 CGGTAGAGGTTCTTTGGCTACACTACTCTCCACCAACGCTGCGCTTCACATCATCTCCGA 743

```











OY 836 99 837  
DB 1350 GG 1351

## RESULT 13

US-08-446-922-5/C  
; Sequence 5, Application US/08446922  
; Patent No. 5716805  
; GENERAL INFORMATION:  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Srinivasan, Subhashini  
; TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric  
; NUMBER OF INVENTION: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.1  
; SOFTWARE: Microsoft Word for Apple, Version 5.1a  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,922  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 08/107,353  
; FILING DATE: 08-13-93  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia A  
; REGISTRATION NUMBER: 34,693  
; REFERENCE/DOCKET NUMBER: 1003-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206)587-0430  
; TELEFAX: (206)233-0644  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 783 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Mouse  
; STRAIN: CD40-L  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..780  
; US-08-446-922-5

Query Match 1.6%; Score 34.2; DB 1; Length 783;  
Best Local Similarity 43.6%; Pred. No. 1.8;  
Matches 153; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

OY 1358 gagggcggaactggagcccgcaaacccactgatacgaacttcaactcgccgt 1417  
DB 774 GAGTAAGCCAAAAGATGAGAGACCAACTCTGTGATCACTTGCTTCAATCAGTCA 715  
OY 1418 gctaatgacagtgagcaatacaagaggttacctctcgatgagtgagaaagctgtgctc 1477  
DB 714 GACAAACACAGAGACCAACGCTTGTATTCACAAACATCCGCCAAGTGAACGACTGCTG 655  
OY 1478 cgcaatgtgcccgcgacactgttgagatcttgcgacttgaagaacaactccaaggttgg 1537

DB 654 CTCGCAAAAGCTGGAGGAGACTGTGGTATTTCGCCCTTGAGTAAGATTCTCTCATGCC 595  
OY 1538 actcaaccctgttcacattcaactctgttgaactccagagtccttccttcagtcacgtcgccgt 1597  
DB 594 ACTGCTGGGCTTCAAGCCAGACAGCCGACGATGAATGGGCTTGACTGGAAGGCTCCGATTT 535  
OY 1598 gagtcgagccttatgagagctgtctgtctcaagatgtgtctgctgctgctgctgag 1657  
DB 534 AGAGCAGAGAGTGACTGTGATGTAGACATAATAGAGTCTTCTCTTTTACCGTCAAGCTG 475  
OY 1658 gtgtctatgttgagggccactaagctccttcctccgtaagttcgccttt 1708  
DB 474 TTTCATTTTCAAGCATTAACCAAGTTGCTTTTCATGTATTAATATCTTT 424

## RESULT 14

US-08-249-189-1/C  
; Sequence 1, Application US/08249189  
; Patent No. 5961974  
; GENERAL INFORMATION:  
; APPLICANT: ARMITAGE, RICHARD  
; APPLICANT: PANSLOW, WILLIAM  
; APPLICANT: SPRIGGS, MELANIE  
; APPLICANT: SRINIVASAN, SUBHASHINI  
; APPLICANT: GIBSON, MARYLOU  
; TITLE OF INVENTION: NOVEL CYTOKINE  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMUNEX CORPORATION  
; STREET: 51 UNIVERSITY STREET  
; CITY: SEATTLE  
; STATE: WASHINGTON  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.1  
; SOFTWARE: Microsoft Word for Apple, Version 5.1a  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/249,189  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/969,703  
; FILING DATE: October 23, 1992  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/805,723  
; FILING DATE: December 5, 1991  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/783,707  
; FILING DATE: October 25, 1991  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia A.  
; REGISTRATION NUMBER: 34,693  
; REFERENCE/DOCKET NUMBER: 2802-C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 2065870430  
; TELEFAX: 2065870606  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 783 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:

ORGANISM: MOUSE  
IMMEDIATE SOURCE:  
CLONE: CD40-L  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..783  
US-08-249-189-1

Query Match 1.6% Score 34.2; DB 2; Length 783;  
Best Local Similarity 43.6%; Pred. No. 1.8;  
Matches 153; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 1358 gagggcggaactggagaccgcaaccccaactgatacgaacttcaacttcgagcgt 1417  
DB 774 GAGTAAGCAAAAGATGAGAGCCAACTGTGATCACTTGCTGGCTGAGTACAGTT 715  
QY 1418 gctaatgacagtgagcaatcaacggagttactctcggatgctgagaacccgtctc 1477  
DB 714 GACAAACACAGAGACCCAGCTTGTAAATTCAAACACTCCGCCAAGTGAACAGACTGCG 655  
QY 1478 cgcgaatgtcccccgcgacactgtgagatctgagcactgagaacactccaaggttgg 1537  
DB 654 CTCGGAAGCTGGGAGAGAACTGTGGTATTTGCCCTTGAGTAAGATTCTCTCAGATCC 595  
QY 1538 actcaacctgttcaactcaactcgttgaactcgaagtccttccgttccactgcccgt 1597  
DB 594 ACTGCTGGCTTCAGCCAGAGCCGACATGAAAGGCGTTGACTGCAAGGCTCCGATG 535  
QY 1598 ggaatcgagcctatgagagcgtgctgctcaagagatgtgtctgagcgtgctgag 1657  
DB 534 AGAGCAGAGGAGCTTGAAGTGAACATAATAGAGTCTTCTTAAACCGTACGCTG 475  
QY 1658 gtgctatgtgagcccaactcgtccttccgtaagtctcgcctt 1708  
DB 474 TTTCATTTTCAACGATTACCAAGTGTCTTTTCATGTGATATATCTTT 424

RESULT 15  
US-08-484-624A-1/c  
Sequence 1, Application US/08484624A  
Patent No. 5962406  
GENERAL INFORMATION:  
APPLICANT: ARMITAGE, RICHARD  
APPLICANT: FANSLAW, WILLIAM  
APPLICANT: SPRIGGS, MELANIE  
APPLICANT: SRINIVASAN, SUBHASHINI  
APPLICANT: GIBSON, MARYLOU  
APPLICANT: MORRIS, ARVIA E.  
APPLICANT: MCGREW, JEFFERY  
TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMUNEX CORPORATION  
STREET: 51 UNIVERSITY STREET  
CITY: SEATTLE  
STATE: WASHINGTON  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,624A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/477,733  
FILING DATE: June 07, 1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/969,703  
FILING DATE: October 23, 1992  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/805,723  
FILING DATE: December 5, 1991  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/783,707  
FILING DATE: October 25, 1991  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia A.  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 2802-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 2065870430  
TELEFAX: 2065870606  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 783 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: MOUSE  
IMMEDIATE SOURCE:  
CLONE: CD40-L  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..783  
US-08-484-624A-1

Query Match 1.6% Score 34.2; DB 2; Length 783;  
Best Local Similarity 43.6%; Pred. No. 1.8;  
Matches 153; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 1358 gagggcggaactggagaccgcaaccccaactgatacgaacttcaacttcgagcgt 1417  
DB 774 GAGTAAGCAAAAGATGAGAGCCAACTGTGATCACTTGCTGGCTGAGTACAGTT 715  
QY 1418 gctaatgacagtgagcaatcaacggagttactctcggatgctgagaacccgtctc 1477  
DB 714 GACAAACACAGAGACCCAGCTTGTAAATTCAAACACTCCGCCAAGTGAACAGACTGCG 655  
QY 1478 cgcgaatgtcccccgcgacactgtgagatctgagcacttgaagaactccaaggttgg 1537  
DB 654 CTCGGAAGCTGGGAGAGAACTGTGGTATTTGCCCTTGAGTAAGATTCTCTCAGATCC 595  
QY 1538 actcaacctgttcaactcaactcgttgaactcgaagtccttccgttccactgcccgt 1597  
DB 594 ACTGCTGGCTTCAGCCAGAGCCGACATGAAAGGCGTTGACTGCAAGGCTCCGATG 535  
QY 1598 ggaatcgagcctatgagagcgtgctgctcaagagatgtgtctgagcgtgctgag 1657  
DB 534 AGAGCAGAGGAGCTTGAAGTGAACATAATAGAGTCTTCTTAAACCGTACGCTG 475  
QY 1658 gtgctatgtgagcccaactcgtccttccgtaagtctcgcctt 1708  
DB 474 TTTCATTTTCAACGATTACCAAGTGTCTTTTCATGTGATATATCTTT 424

Search completed: October 5, 2002, 20:20:55  
Job time: 8530 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:56 ; Search time 19.26 Seconds

(Without alignments)  
1172.040 Million cell updates/sec

Title: US-09-656-640A-2

Perfect score: 3114

Sequence: 1 MISQAGAVAGLAVIGGSS.....IQEMAFNPYAQADDAAE 583

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2010	64.5	572	1	BLRO_MYRVE
2	731.5	23.5	513	1	COTR_BACSU
3	536.5	17.2	642	1	PHSA_STRAT
4	412.5	13.2	516	1	YACK_ECOLI
5	319	10.2	470	1	SUFI_ECOLI
6	310	10.0	470	1	SUFI_SALTY
7	247.5	7.9	622	1	FETS_YEAST
8	247	7.9	311	1	SUFI_HAEPN
9	219.5	7.0	591	1	LACT_CRYPA
10	218.5	7.0	624	1	FETS_CANAL
11	199.5	6.4	527	1	LACS_TRAVE
12	198	6.4	531	1	LACS_THACU
13	196.5	6.3	622	1	YAK8_SCHPO
14	195.5	6.3	520	1	LACT_TRAVI
15	193	6.2	529	1	LACT_PLEOS
16	192	6.2	619	1	LACT_NEUCR
17	190	6.1	636	1	FETS_YEAST
18	188.5	6.1	533	1	LACT_PLEOS
19	188	6.0	548	1	LACT_PHIRA
20	188	6.0	572	1	LACS_THACU
21	187.5	6.0	527	1	LACS_TRAVI
22	187.5	6.0	621	1	LACT_PODAN
23	186.5	6.0	520	1	LACT_TRAVE
24	185.5	6.0	486	1	LACT_BOTCI
25	185.5	5.9	520	1	LACT_TRAVI
26	183	5.9	619	1	LACT_NEUCR
27	182.5	5.9	587	1	ASO_CUCSA
28	178.5	5.7	520	1	LACT_CORNI
29	178	5.7	609	1	COPA_PSESM
30	177	5.7	519	1	LACT_TRAVI
31	176.5	5.7	576	1	LACT_THACU
32	174	5.6	579	1	ASO_CUCMA
33	173.5	5.6	599	1	LACT_THACU

## ALIGNMENTS

RESULT	1	BLRO_MYRVE	STANDARD	PRT	572 AA.
AC	012737				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Bilirubin oxidase precursor (EC 1.3.3.5).				
OS	Myrothecium verrucaria.				
OC	Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Myrothecium.				
OX	NCBI_TaxID=5532;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	STRAIN=MT-1;				
RX	MEDLINE=93366794; Pubmed=8360171;				
RA	Kolkeda S., Ando K., Kaji H., Inoue T., Murao S., Takeuchi K., Samejima T.;				
RT	"Molecular cloning of the gene for bilirubin oxidase from Myrothecium verrucaria and its expression in yeast.";				
RL	J. Biol. Chem. 268:18801-18809(1993).				
CC	- FUNCTION: OXIDATION OF BILIRUBIN AND OTHER TETRAPYRROLES.				
CC	- CATALYTIC ACTIVITY: Bilirubin + O(2) = biliverdin + H(2)O.				
CC	- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR. CONTAINS 2 BLUE COPPER ATOMS PER MOLECULE.				
CC	- SIMILARITY: BELONGS TO THE FAMILY OF PLASTOCYANIN-LIKE DOMAINS.				
CC	- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
DR	EMBL; D14081; BAA03166.1; -				
DR	EMBL; D12579; BAA02123.1; -				
DR	InterPro: IPR001117; Cu-oxidase.				
DR	Pfam: PF00394; Cu-oxidase.1.				
KW	Signal; Copper; Metal-binding; Oxidoreductase; Glycoprotein; Repeat.				
FT	SIGNAL	1	19		
FT	PROPEP	20	38		
FT	CHAIN	39	572		
FT	DOMAIN	98	194		
FT	DOMAIN	404	526		
FT	METAL	132	132		
FT	METAL	134	134		
FT	METAL	172	172		
FT	METAL	174	174		
FT	METAL	436	436		
FT	METAL	439	439		
FT	METAL	441	441		
FT	METAL	494	494		
FT	METAL	495	495		
FT	METAL	495	495		

FT METAL 496 496 COPPER (TYPE 3) (BY SIMILARITY).  
 FT METAL 500 500 COPPER (TYPE 1) (BY SIMILARITY).  
 FT METAL 503 503 COPPER (TYPE 1) (BY SIMILARITY).  
 FT CARBOHYD 510 510 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 520 520 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 572 AA: 63947 MW: 58426413035E5EFF CRC64;

Query Match 64.5%; Score 2010; DB 1; Length 572;  
 Best Local Similarity 65.4%; Pred. No. 6.8e-135;  
 Matches 393; Conservative 71; Mismatches 110; Indels 22; Gaps 10;

QY 1 MISQAIGALVAVIGSSVDARVAGRSTDMPSG-LTKRQTLQSPPLALVLEVPPIPP 59  
 1 MFKHTLGAALSL-LFNSNAVQASVPV--ETSPATGLHFRVQAISPQYMFVPLPIPP 57  
 QY 60 LKAPN-TVPNPNNGEDILYEMETIRPSHOITVDLEPANVGDGMSPGTITVPRKTES 118  
 58 VKQPRLVTPNVPNGOETWYVEIKPETHQVYDLSADLVGDGMSPGTIPVPRVET 117  
 QY 119 VYFVNSGENTSPNSVHLHGSFRAFPDGAEDTQGEYKDYVYPRQARMLWYDHA 178  
 118 VYFVFINNAE--APNSVHLHGSFRAFPDGAEDTQGEYKDYVYPRQARMLWYDHA 175  
 QY 179 MSTANAYWGAQGYMIDDPADALNLPSCYGEEDIPLVYAKRYNADGTLFSTNGEVS 228  
 176 MHTAENAYWGAQGYMIDDPADALNLPSCYGEEDIPLVYAKRYNADGTLFSTNGEVS 225  
 QY 239 SFMGDIYVNGGQWPMPLNVPKRYRFRFLNAVSRSPALVATSEDETLPPQVIAADG 298  
 236 SFMGDIYVNGGQWPMPLNVPKRYRFRFLNAVSRSPALVATSEDETLPPQVIAADG 295  
 QY 299 GLEGEVDVDTLYISMAERWEVYIDSTFAGQSIDIRNLPGA-DGLVEPEFNTDKVMR 357  
 296 GLEGEVDVDTLYISMAERWEVYIDSTFAGQSIDIRNLPGA-DGLVEPEFNTDKVMR 355  
 QY 358 FVYDVELESDETVANLRDVPFEGGNNDPANDD--ETTFEGANQMTINVTFS 415  
 356 FVYADDTPTDTPISVVPANLRDVPF-----SPTTNPFRQRFRTGTGTWTVGVAFA 407  
 QY 416 DVENRLRNVPRTVETWIRLENNNGMTHPVHILVDVRLVSRST--AGVPEYEAAGL 472  
 408 DVONRLRNVPRTVETWIRLENNNGMTHPVHILVDVRLVSRST--AGVPEYEAAGL 466  
 QY 473 KDVYMLARREVVYEAHYAPFPGVYMLHCHNLIHEDHMAAFNVYLGQGYNTEFID 532  
 467 KDVYMLARREVVYEAHYAPFPGVYMLHCHNLIHEDHMAAFNVYLGQGYNTEFID 526  
 QY 533 PMEPLRPRPFLIGFENGSGDFSELATDRIDEMASFNRYAADD 578  
 527 PMEPLRPRPFLIGFENGSGDFSELATDRIDEMASFNRYAADD 572

RESULT 2  
 COTA\_BACSU STANDARD: PRT: 513 AA.

ID COTA\_BACSU STANDARD: PRT: 513 AA.  
 AC P07788; 024818; (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Spore coat protein A.  
 GN COTA OR PIG.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97124186; PubMed=8969499;  
 RA Borris R., Portollik S., Schroeter R.;  
 RT "The 52 degrees-55 degrees segment of the Bacillus subtilis  
 chromosome: a region devoted to purine uptake and metabolism, and

RT containing the genes cota, gabp and guaA and the pur gene cluster  
 RT within a 34960 bp nucleotide sequence.";  
 RL Microbiology 142:3027-3031(1996).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=168 / MARBURG;  
 RX MEDLINE=98116660; PubMed=9455482;  
 RA Kasahara Y., Nakai S., Ogasawara N., Yata K., Sadaie Y.;  
 RT "Sequence analysis of the groESL-cota region of the Bacillus subtilis  
 genome, containing the restriction/modification system genes.";  
 RL DNA Res. 4:335-339(1997).  
 RN [3]  
 RN SEQUENCE OF 1-37 FROM N.A.  
 RX MEDLINE=88286730; PubMed=2821284;  
 RA Donovan W., Zheng L., Sandman K., Losick R.;  
 RT "Genes encoding spore coat polypeptides from Bacillus subtilis.";  
 RL J. Mol. Biol. 196:11-10(1987).  
 RN [4]  
 RN SEQUENCE OF 1-34 FROM N.A.  
 RX MEDLINE=88286730; PubMed=3135411;  
 RA Sandman K., Kroos L., Cutting S.M., Youngman P., Losick R.;  
 RT "Identification of the promoter for a spore coat protein gene in  
 Bacillus subtilis and studies on the regulation of its induction at a  
 late stage of sporulation.";  
 RL J. Mol. Biol. 200:461-473(1988).  
 RN [5]  
 RN SEQUENCE OF 1-10 FROM N.A.  
 RC STRAIN=168;  
 RA Wray L.V., Fisher S.H.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN BROWN PIGMENTATION DURING SPOROGENESIS.  
 CC -1- SIMILARITY: TO S.ANTIBIOTICUS PHENOXAZINONE SYNTHASE (PHSN).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: U51115; AAB62305.1; -  
 DR EMBL: AB007638; BAA22774.1; ALT\_INT.  
 DR EMBL: Z59107; CAB12449.1; -  
 DR EMBL: X05678; CAA29165.1; ALT\_INIT.  
 DR EMBL: X07512; CAA30392.1; -  
 DR EMBL: U31756; AAC44642.1; -  
 DR PIR: A27393; A27393.  
 DR Subtilisin, BG10490; COTA.  
 DR InterPro: IPR001117; Cu-oxidase.  
 DR Pfam: PF00394; Cu-oxidase; 1.  
 KW Sporulation; Complete proteome.  
 FT CONFLICT 347 367 DESRRKPYLASVPSVOHERIQ->TKASRSTSPHTLRS  
 FT CONFLICT 414 420 PRGRTP->RHAETL (IN REF. 1).  
 FT CONFLICT 451 458 GPVAPPPP->VRCRAA (IN REF. 1).  
 SQ SEQUENCE 513 AA: 58499 MW: 8368838458075F87 CRC64;

Query Match 23.5%; Score 731.5; DB 1; Length 513;  
 Best Local Similarity 34.0%; Pred. No. 1.8e-44;  
 Matches 180; Conservative 78; Mismatches 178; Indels 93; Gaps 16;

QY 57 IPLKAPNTPVNPNTGEDILYEMETIRPSHOITVDLEPANVGDGMSPGTITVPRGT 116  
 7 VDALPPIPTLAKPYQSKREKTYEVYMECHTQHLRDLPRRLMGVGLPFGPTIEVKRNE 66  
 QY 117 ESVAFRVNSGENTS-----PNSVHLHGSFRAFPDGAEE-----D 151  
 67 NVYVKMMNNLPSTHPLPDHTIHSDSQHEPEVKTVVHLHGVGTPDDSDGYEAMFSKD 126  
 QY 152 TTQPGEE--KDYTYPRQARMLWYDHAASITANAYWGAQGYMIDDPADALNLPDS 208  
 152 TTQPGEE--KDYTYPRQARMLWYDHAASITANAYWGAQGYMIDDPADALNLPDS 208

```

Db 127 FEQGPYKREYVHPNQQRGAILMWDHMAALRLWVYAGIVGAYIIHDPKERRKLPS 186
QY 209 GYGEDIPLVLTAKRYNADGLETNGE-----VSSFWGDVIOVNGQPMDFLN 256
Db 187 -DEVDLILITDRTIEDGGFLPSPAPENSPSLPMPISIVPACGETIILNKWMPYLE 244
QY 257 VOPKRYRFRFLNAVSRSFALYLTSDSETRLPFOVIADGGLEBPVDTOLYISMAE 316
Db 245 VEPKRYFRVYNASNTKRYNLSTLNGSD----FIOISDGGLLPYSVLNFSFLAPAE 298
QY 317 RMEVVIDIETFAGOSIDIRNLPGADGAGVEEPNDTKVRFVYDELESPTDSEVANL 376
Db 299 RYDIIIDETAEGSIILANAGCGG-DVNPEND-ANIMQFRYTKRLAOKDESRKPYL 355
QY 377 RDVPPEGGNMDFPANP-----TDDETTFGRA-----NGQMTINGVTFSDVENRLRNAP 426
Db 356 ASYPSVOHERIQNTRLTKLAGTOE---YGRPVLLNNKRM-----HDPYTERP 401
QY 427 R-DIVEIMRLENNENGMTHPHILHLVDFRVLRS---STAGVF-----PYEA 469
Db 402 KVGTEETWISIIINPTRG-THPIHLVLSFRYLDRRPFDIARYOESGELSTYGPVPPPSE 460
QY 470 AGLDQVWLAREVYVVEAHAPRPGYUMLHCHLILHEDDDMAAFNVT 518
Db 461 KGMKDTIOAHAGEVLRLAATFGPISGRIYVWCHLLEHEDDDMKRPMDIT 509

```

```

RESULT      3
P_HSA_STRAT
ID   P_HSA_STRAT      STANDARD:      PRT:      642 AA.
AC   053692;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Phenoxazinone synthase (EC 1.-.-.-) (PFS).
GN   P_HSA.
OS   Streptomyces antibioticus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinomycetaceae;
OC   Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxId=1890;
[1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RC   STRAIN=IMKU 5720;
RX   MEDLINE=96011355; PubMed=7592317;
RA   Hsieh C.-J., Jones G.H.;
RT   "Nucleotide sequence, transcriptional analysis, and glucose
RT   regulation of the phenoxazinone synthase gene (phsa) from
RT   Streptomyces antibioticus."
RL   J. Bacteriol. 177:5740-5747(1995).
[2]
RP   SUBUNITS.
RX   MEDLINE=82066838; PubMed=7305384;
RA   Choy H.A., Jones G.H.;
RT   "Phenoxazinone synthase from Streptomyces antibioticus: purification
RT   of the large and small enzyme forms."
RL   Arch. Biochem. Biophys. 211:55-65(1981).
CC   -1- FUNCTION: CATALYZES THE LAST BUT TWO STEPS IN THE PUTATIVE
CC   BIOSYNTHETIC PATHWAY OF ACTINOMYCIN
CC   -1- CATALYTIC ACTIVITY: 4 4-methyl-3-hydroxyanthraniloyl pentapeptide
CC   + 3 O(2) -> 2 actinomycinic acid + 6 H(2)O.
CC   -1- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH
CC   CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE
CC   2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC   -1- SUBUNIT: HOMODIMER (SMALL FORM) OR HOMOHEXAMER (LARGE FORM).
CC   -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC   -1- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.
-----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC      or send an email to license@lsb.sib.ch).
DR      EMBL, U04283; AAA86668.1; ALT_INIT.
DR      InterPro; IPR001117; Cu-oxidase.
DR      InterPro; IPR002355; MulticCu_oxidase.
DR      Pfam; PF00394; Cu-oxidase_1
DR      PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
DR      PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW      Oxidoreductase; Repeat; Metal-binding; Copper;
KW      Antibiotic biosynthesis.
FT      INT_MET      0      0
FT      DOMAIN      86      223      PLASTOCYANIN-LIKE 1.
FT      DOMAIN      493      620      PLASTOCYANIN-LIKE 2.
FT      METAL      161      161      COPPER (TYPE 2) (BY SIMILARITY)
FT      METAL      163      163      COPPER (TYPE 3) (BY SIMILARITY)
FT      METAL      201      201      COPPER (TYPE 3) (BY SIMILARITY)
FT      METAL      203      203      COPPER (TYPE 3) (BY SIMILARITY)
FT      METAL      524      524      COPPER (TYPE 1) (BY SIMILARITY)
FT      METAL      527      527      COPPER (TYPE 1) (BY SIMILARITY)
FT      METAL      529      529      COPPER (TYPE 2) (BY SIMILARITY)
FT      METAL      602      602      COPPER (TYPE 3) (BY SIMILARITY)
FT      METAL      603      603      COPPER (TYPE 3) (BY SIMILARITY)
FT      METAL      604      604      COPPER (TYPE 1) (BY SIMILARITY)
FT      METAL      608      608      COPPER (TYPE 3) (BY SIMILARITY)
FT      METAL      613      613      COPPER (TYPE 1) (BY SIMILARITY)
SQ      SEQUENCE      642 AA; 70113 MW; E0B39CDBA3364E48 CRC64;

```

Query Match	17.2%;	Score 536.5;	DB 1;	Length 642;
Best Local Similarity	29.5%;	Pred. No. 1.6e-30;		
Matches 184;	Conservative 65;	Mismatches 200;	Indels 175;	Gaps 26

QY	37	LTROQOLSPR--LALXEVPLRPPRLAPNTVPNPNGEDILLYEMELRPPSHOITYDLE	94
Db	28	LAKEEQAPAGELTPRPAFLTPVP-----VLBPDADEVTREREIALRPTWMLHOLP	81
QY	95	PANNVYDGMSPGPTIIV-----PRGTESVRFV-----NGSENTSP-131	
Db	82	PTLMWVGDOVPGPTIEVRGQRARIAMTNKIPRGSEYPTVSVPRLPGCTAPRNPTEPG	144
QY	132	-----NSVHLHGSFSRAPRDEGMAEDTTQGEYKDDYYYPNRQAAAMLWYH	175
Db	142	RGVEVPNKDVAAALPMAWSTYHLHGQOTGGGNDGMADNVAVGDAQDLSXPMDHQATQWYH	201
QY	176	DHAMSTIAENMYMOAGVYMMIODEAEALNIPSGTEGEDIPIVYLAKRYAAD-----GT	229
Db	202	DHAMITTRMWNMACLYGTLYLRDDEDEALGPLS--GIREIPLTLADNNLDTDEGRUNGR	259
QY	230	LF-----STNGCVSS-----FWGADVIOVNGQOPWPMILNPOKRKYRFRFLNAAVSRSAL	277
Db	260	LLHKTVLYQGSNPRTGKRPVSIPIFGGPTTYTNGRIWPRADVDGCGRIURLVLYNAAKRIYNL	319
QY	278	YLAFTSESEFRLPRQVYLAADGGELGFEVTD-----TLYISMAERWVVIDPSTFAGOS	331
Db	320	VLIDEDDRPVPVGVYHQIGSGGLLRPVPVDFDTEPLVLSAAAPERDILLVDFRALLGRR	379
QY	332	IDIRNL-PGA-----DGLG--VBEERPNITKVMRFVYUDEVLESDDTSEVPAHL-----	376
Db	380	LRLVDKRGAGACPTPDLGCVRYPE-----VMEFRKRETCSE--DSFALPEVLYSGSFRM	432
QY	377	RDVPF-----PEGGN--W-----DPAN-----PTDDEFTTFG	401
Db	433	SHDIPGHRLILVLRPGTKSGGGRPELTWMAEVEDPADVOVPAAGVLYQVGAAGRITTYR	422
QY	402	RANQMTINGVTSDEVNRLRLNRPDUTVEIWRLENNSGMTHPEVHILHDFRVLXSTA	461
Db	493	R-----TATFND-----GLGFTIGEGTGHQWAFNLNS--PILHPMIIHLADFQVLGR--	538
QY	462	RGVEPYEAGL-----KDYVWILARREYVYEAHYAPR-----	494
Db	539	---DATASGEFDLALGSTRIPVRLDPDTPVPLAPNELNGHADVQVDFEPOGLRWMGKFQDA	595
QY	495	-GVYMLHCHNLIHEDHDMAAFNV	517

DB 596 YGRFMYHCHLHEHDMGMRRPFV 619

RESULT 4

YACK\_ECOLI STANDARD: PRT: 516 AA.

AC P36649; P75655;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable blue-copper protein yack precursor.

GN YACK OR B0123.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_TaxId=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / W3110;

RX MEDLINE-94261430; PubMed-8202364;

RA Fujita N., Mori H., Yura T., Ishihama A.;

RA the 2.4-4.1 min (110,917-193,643 bp) region.";

RT Nucleic Acids Res. 22:1637-1639(1994).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RX MEDLINE-97426617; PubMed-9278503;

RA Blatner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.C., Rose D.J.,

RA Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP SEQUENCE OF 29-40.

RC STRAIN-K12 / EMG2.

RX MEDLINE-97443375; PubMed-9298646;

RA Link A.J., Robinson K., Church G.M.;

RA "Comparing the predicted and observed properties of proteins encoded

RT in the genome of Escherichia coli K-12.";

RL Electrophoresis 18:1259-1313(1997).

RN [4]

RP IDENTIFICATION BY MASS SPECTROMETRY.

RX MEDLINE-99420866; PubMed-10493123;

RA Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;

RA "Enrichment of low abundance proteins of Escherichia coli by

RT hydroxyapatite chromatography.";

RL Electrophoresis 20:2181-2195(1999).

CC -1- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH

CC CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE

CC 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: periplasmic (potential).

CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.

CC -1- SIMILARITY: CONTRAINS 3 PLASTOCYANIN-LIKE DOMAINS.

CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A

CC FRAMESHIFT IN POSITION 464.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL: D26562; CAB20297.1; ALT\_FRAME.

DR EMBL: AE000121; AAC73234.1; -.

DR PIR: S45200; S45200.

DR Ecogene: Egi2318; yack.

DR InterPro: IPR001117; Cu-oxidase.

DR InterPro: IPR002355; MulticCu\_oxidase2.

DR Pfam: PF00394; Cu-oxidase; 1.

DR PROSITE: PS00080; MULTICOPPER\_OXIDASE2; 1.

KW Periplasmic; Signal; Copper; Metal-binding; Repeat; Oxidoreductase;

KW Complete proteome.

FT SIGNAL 1 28

FT CHAIN 29 516 PROBABLE BLUE-COPPER PROTEIN YACK.

FT DOMAIN 67 163 PLASTOCYANIN-LIKE 1.

FT DOMAIN 164 410 PLASTOCYANIN-LIKE 2.

FT DOMAIN 411 516 PLASTOCYANIN-LIKE 3.

FT METAL 101 101 COPPER (TYPE 2) (BY SIMILARITY).

FT METAL 103 103 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 141 141 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 143 143 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 443 443 COPPER (TYPE 2) (BY SIMILARITY).

FT METAL 446 446 COPPER (TYPE 2) (BY SIMILARITY).

FT METAL 448 448 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 489 489 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 500 500 COPPER (TYPE 1) (BY SIMILARITY).

FT METAL 501 501 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 505 505 COPPER (TYPE 1) (BY SIMILARITY).

FT METAL 510 510 COPPER (TYPE 1) (BY SIMILARITY).

SO SEQUENCE 516 AA; 56556 MW; 37D96B1C31CF30B CRC64;

Query Match 13.2%; Score 412.5; DB 1; Length 516;

Best Local Similarity 29.7%; Pred. No. 6,9e-22;

Matches 141; Conservative 55; Mismatches 196; Indels 83; Gaps 19;

QY 100 GYGVGSGPGPIIYVPRGTESVVRVNSGEMTSPNSVHLSFSAPRPDGMADDTTQGEYK 159

DB 68 GINGNLGPVVKIQKRGKAVTDIYN--QLTEETLLMHGLVPEVDGGPQGITPPGKR 125

QY 160 DYVYPRROARMLMYHDHMSITAEVNAVMGQGVVMIODPAEDALNLPSCGEFDIPLV 219

DB 126 SVTLNDOPATQCFWPHDHGKTGRQVAMGLAVIEDEDLIKLMPQMGIDVPIV 185

QY 220 TAKRYADGTL---ESTNGEVSSFMGDIQVNGQPRMLNVQPRKI--RRFLNAVSRSF 275

DB 186 QDKKFSADQIDYQLDVMTAAVGMFGDITLTNGAIPQ--HAARGLRLRLNGCNARS- 243

QY 276 ALVLATSESETRLPVOVIAADGGLLEGVPDITLISMAEREVVIDESTFAGQSIDIR 335

DB 244 -LNFATSDNR---PLVYIASDGLLPPEVYKSELVYLNGEREVLVEVND--NKFFDLV 296

QY 336 NLPGAD-GLGVBEFDNTKVMRFVDEVLES---PDT-SEVPANLADVPEFG----- 384

DB 297 TLPSQMGAIAP-FDKPRVPMIQTPLAISAGALPDTLSIPA---LPSLEGLTVRK 351

QY 385 -----GNMDPAPTDEFTTFRANG 405

DB 352 QLSMDMLDMGMQMLMERYGDQAMAGMDHSQMGHMGNNMHN--HGKFDFFHAN- 408

QY 406 QWTLNVFSDVENRLRLRVPRDTVEIWRLENNNSNGWTPTVNHLDVFLVSTARGVE 465

DB 409 --KINGQARD--MKNKMFPAAKQGYERWYISGVGMMLRPFHNGIQFRLISN---GAP 461

QY 466 P-YEAGLDVYVLARR--EYVYVNAHYAPFPFGVYVLMHCHNLHEDHMAAFENV 517

DB 462 RAHNRAGWKDYVXVEGVNSEVLVKFNHDAPKEHAYVMAHCHLLEHEDTGMMLGFTV 516

RESULT 5

SUFL\_ECOLI STANDARD: PRT: 470 AA.

ID SUFL\_ECOLI

AC P26648;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Protein sufl precursor.

GN SUFL OR B3017.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

```

OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=92278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE OF 1-89 FROM N.A.
RX MEDLINE=92212294; PubMed=1557036;
RA Coleman J.;
RT "Characterization of the Escherichia coli gene for
RL 1-acyl-sn-glycerol-3-phosphate acyltransferase (plsc).";
RL Mol. Gen. Genet. 232:295-303(1992).
RN [3]
RP SEQUENCE OF 28-39.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RL in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1329-1313(1997).
CC -1- FUNCTION: INVOLVED IN CELL DIVISION, SUPPRESSES A FTSI MUTATION.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
-----
DR EMBL: U28377; AAA69185.1;
DR EMBL: AE000384; AAC76053.1;
DR EMBL: M63491; AAA24398.1;
DR PIR: S20461; S20461.
DR EcoGene: EG11376; SUFI.
DR InterPro: IPR001117; Cu-oxidase.
DR Pfam: PF00394; Cu-oxidase; 1.
DR KEGG: Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 27
FT CHAIN 28 470 PROTEIN SUFI.
FT DOMAIN 68 164 PLASTOCYANIN-LIKE.
FT SEQUENCE 470 AA; 51858 MW; C843A54ACB146688 CRC64;
SQ

```

Query Match 10.2%; Score 319; DB 1; Length 470;  
 Best Local Similarity 25.3%; Pred. NO.2.5e-15;  
 Matches 131; Conservative 67; Mismatches 205; Indels 114; Gaps 21;

```

QY 10 ALGLAVIGSSVDARSVAGRSTDMPSGLTKRQTLSPPLALYEYPLPIPLKAPNTPNP 69
DB 11 ASGIALCAGAVPLKASNAAG-----QQPLPVPLLL-----ES-42
QY 70 NTEEDILYEMETRPFSHOITPLEPANNVGYQMSGPIITVPRGTESVYRVNSENTE 129
DB 43 RKQDFE--MTYQR-AHMSFTPGTRASVINGIRYLGPIRWKGD--VKLYSNRLT 96
QY 130 SPNSVHLGHSFSAPEFGMAEDTTPGCEYKDYYPNROAARMLWYHDHAMSITAENYMG 189
DB 97 ENVSMTYAGIQVCEPLMGGRARMSAPADAPVLPIDONATITMYHANTPTNRQAQVYNG 156
QY 190 QAGVYMIQDAEDALNLPSCYGEDEPLVLTAKRYNADGTLFTNGEVSFSGDVQVNG 249
DB 157 LAGMMLVEDEVSKSLPIPNHYGVDDFVITIQDKRLDNFGTRPEVNEPSGSGFVDTLLVNG 216
QY 250 QPAPMLNVQPRKTRFRFLNAVSRSPALYLATSEDETRLPFOVIADGGLLEGVDITPT 309

```

```

DB 217 VQSPYVEVSRGWRLRLNLSNSRRYQLOMNDGR-----PLHVISQDQFLPAPVSKQ 270
QY 310 LYISMARREYVIDFST-----FAGQSDIDINL-----PGADGL 343
DB 271 LSLAPGERRELVDMSNGDEVSTICGEASIVDRINGCFERSSILVSTLYLNP-----T 326
QY 344 GVEPEFNTKRV-NRFVAVDEVLESPTSEVPANLRDVPPEGGWMDPANPTDDETFEGR 402
DB 327 GLPLVY--TDSLPRLLPTEIMAGS-----PISRDISLGD---DYG----- 363
QY 403 ANQO-WTINGVTSDEVNRLLRNVRDTEIWRLENNNSNGTHPHTHLYDFRYSRSTA 461
DB 364 INGLMDVNRK-----DY-----TAQGTWERWYTRADE---PQAFHLEGVWFQI---RVN 408
QY 462 RGVPEY-EEAGLADVVL---AREVYVEAHVAFPP 494
DB 409 NGAMPFEDRGWKDTVWVDQVELLYVFGQPSMAHFP 445

```

RESULT 6  
 SUFI\_SALTY STANDARD; PRT; 470 AA.  
 ID SUFI\_SALTY  
 AC P40799;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Protein sufi precursor  
 GN SUFI OR STM3172 OR STY3349.  
 OS Salmonella typhimurium, and  
 OS Salmonella typh.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=602, 601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhimurium; STRAIN=LT2;  
 RA Cong J., Schmid M.B.;  
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SSGC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RL LT2.";  
 RL Nature 413:852-856(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhimurium; STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,  
 RA Churcher C., Mangell K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Barrar J.,  
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,  
 RA Krogli A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrall B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 RT enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 CC -1- FUNCTION: INVOLVED IN CELL DIVISION, SUPPRESSES A FTSI MUTATION  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).  
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 413.



FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 622 AA; 70880 MW; 961757A6C86B7FAF CRC64;

Query Match 7.9%; Score 247.5; DB 1; Length 622;  
 Best Local Similarity 22.3%; Pred. No. 4.2e-10;  
 Matches 130; Conservative 88; Mismatches 187; Indels 177; Gaps 31;

OY 91 PD-LEPANNVGDMSBPPTIIVRGTSYVRFVNSGENTSPNSVHLGSF-----SRA 143  
 DB 34 PDLHKKRMJGFWPLPDHVEKSDRYELYLNGFDATATSLRHGLFQNTSLGNOL 93  
 OY 144 PFGMAEDTQ-----PGEYDYYPNQAARMLYHDHNASITAEANAKQAQAYMTQDP 199  
 DB 94 QMDG-PSMTQCPVPSQYLYNFVPEQVGTWYAHMGAQYGD-----GMRGAFIHD 148  
 OY 200 AEDALNLPSSGGEFDIPLVLT-----AKRYNADGT-----LESTNGE 236  
 DB 149 EEPF-----EYDHERVITLSDHYENKTYTKETLSYNTPGAEPITQNTLNNMTAN 200  
 OY 237 VSSEFGDVIQVNGQPMPLNVQ-RRYRREFLNAVSRFALYLAISESETRLPQVIA 295  
 DB 201 VT-----LDFTEGYLFRFLNVGLFVSQYIL-----EDHE-----MSIVE 237  
 OY 296 ADGGLGPDVDTLYISMAERWEVVDPS-----TFA-----GQ 330  
 DB 238 VDSVYVK-PNFTDSTYLSAQARMSVILKAKDKMPTRYAMQIMDTMLDVPPELOLN 296  
 OY 331 SIDIR---NLPGADGLGVEP-EED--NTDKVMPFVDEVPDTESEVPANLRDVPPEG 384  
 DB 297 TIGMRGHSLEPARALNIEBCDDLRANDPYLERLLEROLLAYHDQIYDVAMNLGDS 356  
 OY 385 GNV-----DPANPTDETFEGRANGQWITNGVTSDEVNRLRNVPDTEIWL 435  
 DB 357 VKAFENNTIYVYPKPTLTLLTSGKLASDPRIYG--DINAQLKHN--NDIIEV--V 409  
 OY 436 ENNSNGWTPVHNLVDFRYSRSTARGV-----PY-EAGL-----KDY 436  
 DB 410 LNNYDGRPHFILGHNFQVQKSPGHVDEAYDESEODEMTVPYNESAPLOPPEPMP 469  
 OY 477 WLAREVYVE--AHY-----APFPGVYMLCH-----NLH 506  
 DB 470 ----RDTVVLPSGCHVYLRADNPGWYFCHCHVMDHLOGLASVFIETAVLLQEREKLN 525  
 OY 507 EDH-DMAAARNVTVLGDYGYNTETID---PNEPLMRPRPF 543  
 DB 526 ENYLDICKADIPVVGNAAGHSNDWFDLGLRQPEPLPKGF 567

RESULT 8  
 SUPTL\_HAEIN STANDARD: PRT: 311 AA.  
 AC P44847;  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Protein sufi homolog precursor.  
 GN SUFI OR H10733.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus  
 NCBI\_TaxId=727;  
 OX 11  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Keilavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McInney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 influenzae Rd";  
 RL Science 269:496-512(1995).  
 CC - SUBCELLULAR LOCATION: Periplasmic (Potential).  
 CC - SIMILARITY: STRONG, TO E.COLI AND S.TYPHIMURIUM SUPTL.  
 CC - SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U32756; AAC22390.1; .  
 DR TIGR: H10733; .  
 KW Periplasmic; Signal; Complete proteome.  
 FT SIGNAL 1 28 BY SIMILARITY.  
 FT CHAIN 29 311 PROTEIN SUFI HOMOLOG.  
 FT DOMAIN 68 164 PLASTOCYANIN-LIKE.  
 SQ SEQUENCE 311 AA; 34496 MW; 436173F59F1E1772 CRC64;

Query Match 7.9%; Score 247; DB 1; Length 311;  
 Best Local Similarity 31.1%; Pred. No. 1.8e-10;  
 Matches 96; Conservative 40; Mismatches 129; Indels 44; Gaps 13;

OY 33 MBSGLTKRQ-----TQLS-----PPLALYVPLPPLKAPNTVPNPTGEDILYEM 80  
 DB 1 MPR-LSRQLTKTAISTALSTVAPAPLAAAREKLVPPL-----IEVRGRPIVLTWQ 53  
 OY 81 EIRPFHQIYPOLEPNMV--GYDGMSPGPTIIVRGTSYVRFVNSGENTSPNSVHL- 136  
 DB 54 ETN-----YF-LDGSNNVYWGFGNGVYLGPTIKRISGSFAKLN--HNNLPQVALS 102  
 OY 137 -HGSFRAPEFDGWAEDTQPEYKDYYPNQAARMLYHDHNASITAEANAKQAQAYVM 195  
 DB 103 IQGLQASGELFGGAARVLKKGESMAPIVTEQPAACGWYSATLANSAYQTYGLGML 162  
 OY 196 IQDPAEDALNLPSSGGEFDIPLVLTAKRYNADG-TLESTNGEVSSFGVDVIQVNGQPM 254  
 DB 163 IEDEQSLKALPNKRYGVDDIPLIQDMEFNNDLQLEFKQN--QPHFVGNRLVNGIEAPY 220  
 OY 255 LNVQPKRYRRLNMAVSRFALYLAISESETRLPFOVIAADGGLGPDVDTLYISM 314  
 DB 221 LDVARGWIRLRLNMAVLARAYDLRL--DNDQEMLL---IAQDLGLPRAKYSKSVLSP 274  
 OY 315 AERREVID 323  
 DB 275 GERAEITLVN 283

RESULT 9  
 LAC1\_CRYPA STANDARD: PRT: 591 AA.  
 AC 003966;  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Laccase precursor (EC 1.10.3.2) (Benzeneol: oxygen oxidoreductase)  
 DE (Urishiol oxidase).  
 GN LAC-1.  
 OS Cryphonectria parasitica (Chesnut blight fungus) (Endothia  
 parasitica).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Diaporthales; Valsaceae; Cryphonectria.  
 OX NCBI\_TaxId=5116;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92314427; PubMed=1535523;
RT Chai G.H., Larson T.G., Nuss D.L.:
RT "Molecular analysis of the laccase gene from the chestnut blight
RT fungus and selective suppression of its expression in an isogenic
RT hypovirulent strain."
RL Mol. Plant Microbe Interact. 5:119-128(1992).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzenesemiquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted (potential).
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M73257; AAA3105.1; -
CC EMBL: S38903; AAA09235.1; -
CC HSRP: P37064; IAO2.
CC InterPro: IPR001117; Cu-oxidase.
CC InterPro: IPR002355; Multicu_oxidase2.
CC Pfam: PF00394; Cu-oxidase; 3.
CC PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
CC PROSITE: PS00080; MULTICOPPER_OXIDASE3; 1.
CC Oxidoreductase; Signal; Copper; Metal-binding; Lignin degradation;
CC Glycoprotein; Repeat.
CC KMW Glycoprotein; Repeat.
CC FT SIGNAL 1 20
CC FT CHAIN 21 591
CC FT DOMAIN 66 189
CC FT DOMAIN 198 356
CC FT DOMAIN 416 551
CC FT METAL 126 126
CC FT METAL 128 128
CC FT METAL 171 171
CC FT METAL 173 173
CC FT METAL 463 463
CC FT METAL 466 466
CC FT METAL 468 468
CC FT METAL 533 533
CC FT METAL 534 534
CC FT METAL 535 535
CC FT METAL 539 539
CC FT METAL ? ?
CC FT CARBOHYD 121 121
CC FT CARBOHYD 224 224
CC FT CARBOHYD 242 242
CC FT CARBOHYD 265 265
CC FT CARBOHYD 323 323
CC FT CARBOHYD 407 407
CC FT CARBOHYD 425 425
CC SEQUENCE 591 AA; 64696 MW; B2F44CB2AD77701 CRC64;

```

Query Match 7.0%; Score 219.5; DB 1; Length 591;  
 Best Local Similarity 23.6%; Pred. No. 3.8e-08;  
 Matches 155; Conservative 78; Mismatches 228; Indels 195; Gaps 39;

```

QY 35 SGLTKRQTOISPLALYEVLPPIPKAPN-----TVPNPTGE 73
DB 10 SGLIASQLSMAAPSLTH-----PLEPRQOPNCNTASNRACWISGSYDITTDYEVKPLTGV 65
QY 74 DILYEMETIRPFSHQIYTPD-LERANVGVGDGMSPGPIIIVPRTESEVVRVNGENTSPN 132

```

```

DB 66 -VRQYDLTLTQAEENWLGPPGVVKEEDVLMVNGNLTGVTIAQWG-DTISVTNNLKYNGT 123
QY 133 SVYLIHG-----SFSRAPFPGMAEDTTP--GEYKDYYPQRAARMVLYDHAMSITAENA 186
DB 124 TIHHMCIROLNNTLQGVNGITCEPIPRNGSKTYTFFIAHQVGT- WYSH- FSAQYGN- 180
QY 187 YNGQAGVYMIQDPAEDALNPSQYGEFDI--PLVITAKRY-NADGTLFSTNGEVSFWG 242
DB 181 --GIYGAIQIDGPA-----SLP-----YDIDLPLVLSDDYKYTADLVLVYTSQSNAPPA-S 228
QY 243 DYIVVNGQCPWPHLNVOP-----RKRRFFFLAASRSFALVLAISSEDE 286
DB 229 DNVLFNG-----TINPANTTGGQYKTTTLTPGKRRLRIINTSVANNFOVSIVGHS--- 280
QY 287 TRLPFVIAADGGLLEGPPVD--TDLTYSMAERMEVVDFTS-----FA 328
DB 281 ----MTVIESD-----FVPPDSFTTDSLEFGIGQRYVITIDAQATNTYMMNTYFGGCGFC 332
QY 329 GOSID-----IRNLPGADGLGVEPEFNDTKVREYVDEVLESPP-TSEVPANLRDVPFP 382
DB 333 GKSNNPYPAALIHYNGA-----SNHPNKKGVAPADHECLDLNLVPPVPRSP-TP 382
QY 383 EGNMNPAMPPTDEFTTFRANGQWTINGVTSQVE-----NRLRNVPRTYDEI 432
DB 383 SGFVAASDNTLDVQLSTTR--KWTINSGTL-DVDMGHPITQYIYNKSTAMPSTDNWVL 438
QY 433 -----WRLENN-----SNGWTHPHVHLVDFRILSR-----TARVEPEYEAAGLMD 474
DB 439 VERANQWAWLIENTPTAGNLPHRPHLHGDFVVLGSSPNVSPATQ--TPPTFTS-SD 495
QY 475 VVWL-----ARREYVVE-----AHVAPFQGVYWLHGHNLHEDHDMAAFNVTVLGD 522
DB 496 VSSLLGNNGNPIRRVYVWLPRKGMILLIAFQTTNGGAMLMHCHIMW-----VSAG 543
QY 523 YGVNTEFTDPMPLMRPRPPLIGEFENGSGPSELAITDRQEMASFNPAQAD 578
DB 544 LGNTFLE-----QPSNFAVAGLNTN--DVNQLN-----SQCKSMNAYTPSKD 582

```

```

RESULT 10
FET3_CANAL STANDARD; PRT; 624 AA.
AC F78591;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Iron transport multicopper oxidase precursor (EC 1.-.-.-).
GN FET3.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1161;
RL Hunt S., Eck R.;
RT Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RT -1- FUNCTION: IRON TRANSPORT MULTICOPPER FERROXIDASE REQUIRED FOR
RT FERROUS IRON HIGH AFFINITY UPTAKE. REQUIRED TO OXIDIZE FE(II) AND
RT RELEASE IT FROM THE TRANSPORTER. ESSENTIAL COMPONENT OF COPPER-
RT DEPENDENT IRON TRANSPORT (BY SIMILARITY).
CC -1- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH
CC CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE
CC 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL SURFACE BOUND
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE MULTICOPPER OXIDASE FAMILY.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```



CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Y09329; CAA/0509.1; -  
DR HSSP: P37064; IAOZ.  
DR InterPro: IPR001117; Cu-oxidase.  
DR InterPro: IPR002355; Multicu\_oxidase2.  
DR Pfam: PF00394; Cu-oxidase; 3.  
DR PROSITE: PS00079; MULTICOPPER\_OXIDASE1; 1.  
DR PROSITE: PS00080; MULTICOPPER\_OXIDASE2; 1.  
KM Oxidoreductase; Copper; Iron transport; Transport; Repeat;  
KM Transmembrane; Signal; Glycoprotein.  
FT SIGNAL 1 20 POTENTIAL.  
FT CHAIN 21 624 IRON TRANSPORT MULTICOPPER OXIDASE.  
FT DOMAIN 46 144 PLASTOCYANIN-LIKE 1.  
FT DOMAIN 190 292 PLASTOCYANIN-LIKE 2.  
FT DOMAIN 382 499 PLASTOCYANIN-LIKE 3.  
FT DOMAIN 21 555 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 556 576 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 577 624 POTENTIAL.  
FT METAL 81 81 COPPER (TYPE 2) (BY SIMILARITY).  
FT METAL 83 83 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 126 126 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 128 128 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 413 413 COPPER (TYPE 1) (BY SIMILARITY).  
FT METAL 416 416 COPPER (TYPE 2) (BY SIMILARITY).  
FT METAL 418 418 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 481 481 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 482 482 COPPER (TYPE 1) (BY SIMILARITY).  
FT METAL 483 483 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 487 487 COPPER (TYPE 1) (BY SIMILARITY).  
FT METAL 487 487 COPPER (TYPE 1) (BY SIMILARITY).  
FT CARBOHYD 49 49 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 77 77 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 113 113 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 194 194 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 198 198 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 244 244 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 265 265 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 292 292 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 359 359 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 441 441 N-LINKED (GLCNAC...) (POTENTIAL).  
SO SEQUENCE 624 AA; 70635 MW; EE82B21F32346918 CRC04;

Query Match 7.0%; Score 218.5; DB 1; Length 624;  
Best Local Similarity 24.4%; Pred. No. 4.8e-08;  
Matches 123; Conservative 66; Mismatches 166; Indels 149; Gaps 29;

QY 91 PD-LEPANNVGYDGMSPRTIIVRGTESVYRFPNSGENTSPNSVHLHGSSRA--PEDG 147  
DB 37 PDGVPYPRKMGFNSWDLPLRLVKKGRVOLDYLINGFDNLN-TLHLHGLFVRANQMDG 95  
QY 148 WAEDTTO-----PGEYKXYYPNROAARMVYHDHAMSITANAYAGQVYMIOD----- 198  
DB 96 -PEWVTCPLPPEGETYLYXNFVTDVCGTYWYHSH-----TGQYVDDGAKRGVYIIEEDDPY 150  
QY 199 -PAEDALNLPSSGYEE--DI-PLVLTAKRYNADGT-----LFTNGEVSSEFWGVDIQ 246  
DB 151 HYDEEVVLTSLDHYKYSIGDIPAFLE--RFPNPGABEIPDNFLFN-----E 195  
QY 247 VNGCPMPLMNQPRK-YRFRPLNAVSRFALYIATSEDSSTRLPVOVIADGGLGEPV 305  
DB 196 TRNATW---KVEPKTYFVRLINGVFSOYLWM---EDHE---FTVELDGYVEKNT 245  
QY 306 DTDPLTYISMAERWEVVI-----DFTFAGOSIDIRNLPGADGLGVEPEFNTDKVAREV 359  
DB 246 -TDILITVAVQRYGLVLTTRKSTDKNYFVNGVDITML---DSVPADIQVNGTYI--V 298  
QY 360 VDEYLESFD-----TSEVPANLRDVPPEEGNM 387

DB 299 YNESSALPDVAIDISYDADLDFYLFKPLSKOLMDADATYITVDYQMVNLN-----DGINTY 354  
QY 388 -----DANPTDDEFTTFPGRANGOWTNGYTFPSVEYERLLRNPDRDVEIWRLENN 438  
DB 355 AEFNNISYKARKVPYTLTLVLSAGEATNELTYG---TNTNSFVLGG--GDIVDI--VLNN 407  
QY 439 SNGWTHPVHILVDFPRVLSRSTARGVEPEYEAAGLKDVY-----WLARREVVY 485  
DB 408 FDTGKHHPHLGHVGFQLER-----HEATGSKESAVENTNSDHAEMPEYPIARDIYV 459  
QY 486 VEAH-----YAPFPGVYMLDCH 502  
DB 460 VKPSSHVMYLFREKADNPVYWFHCH 483  
RESULT 11  
LACS\_TRAVE STANDARD; PRT; 527 AA.  
ID LACS\_TRAVE  
AC 012717;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Laccase 5 precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)  
DE (Trisliol oxidase) (Diphenol oxidase) (Laccase IV).  
GN LCC5 OR LCCIV.  
OS Trametes versicolor (White-rot fungus).  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
OC Aphyllophorales; Coriolaceae; Trametes.  
OX NCBI\_TaxID=5325;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=527;  
RX MEDLINE=97464057; PubMed=9322748;  
RA Ong E., Pollock W.B., Smith M.;  
RT "Cloning and sequence analysis of two laccase complementary DNAs from  
the ligninolytic basidiomycete Trametes versicolor."  
RL Gene 196:113-119 (1997).  
-1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED  
PRODUCTS (PROBABLE).  
-1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzosemiquinone + 2  
H(2)O.  
-1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU  
CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE  
3 OR COUPLED BINUCLEAR (BY SIMILARITY).  
-1- SUBCELLULAR LOCATION: Secreted.  
-1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.  
-1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U44431; AAC49829.1; -  
DR HSSP: P37064; IAOZ.  
DR InterPro: IPR001117; Cu-oxidase.  
DR InterPro: IPR002355; Multicu\_oxidase2.  
DR Pfam: PF00394; Cu-oxidase; 3.  
DR PROSITE: PS00079; MULTICOPPER\_OXIDASE1; 1.  
DR PROSITE: PS00080; MULTICOPPER\_OXIDASE2; FALSE\_NEG.  
KM Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;  
KM Lignin degradation; Multigene family.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 527 LACCASE 5.  
FT DOMAIN 25 150 PLASTOCYANIN-LIKE 1.  
FT DOMAIN 162 306 PLASTOCYANIN-LIKE 2.  
FT DOMAIN 373 498 PLASTOCYANIN-LIKE 3.  
FT METAL 87 87 COPPER (TYPE 2) (BY SIMILARITY).  
FT METAL 89 89 COPPER (TYPE 3) (BY SIMILARITY).

Query Match	6.4%;	Score 199.5;	DB 1,	Length 527;
Best Local Similarity	23.3%;	Pred. No. 8.4e-07;		
Matches 120; Conservative	62;	Mismatches 200;	Indels 133;	Gaps 28

```

DE      15-Jul-1999 (Rel., 38, Last annotation update)
DE      Lacasee 4 precursor (EC 1.10.3.2) (benzenediol:oxygen oxidoreductase)
DE      (urushiol oxidase) (Diphenol oxidase).
GN      LCC4.
GN      Thanatephorus cucumeris (Black scurf of potato) (Rhizoctonia solani).
OC      Eukaryota; Fungi; Basidiomycota; Hymenocytetes; Heterobasidiomycetes;
OC      Heterobasidiomycetales; Ceratobasidiales; Ceratobasidiaceae;
OC      mitosporic Ceratobasidiaceae; Rhizoctonia.
CX      NCBI_TaxID=107832;
XX      [1]
RN      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP      STRAIN-RS22;
RC      MEDLINE=961171523; PubMed=8598061;
RA      Wallethner J.A., Xu F., Brown K.M., Brown S.H., Golightly E.J.,
RA      Haliker T., Kauppinen S., Pederson A., Schneider P.;
RT      "Type identification and characterization of four lacases from the
RT      plant pathogenic fungus Rhizoctonia solani.";
RL      Curr. Genet. 29:395-403(1996).
CC      -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC      PRODUCTS (PROBABLY). THIS ISOZYME SHOWS OPTIMAL ACTIVITY AT PH 7.
CC      -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzosemiquinone + 2
CC      H(2)O.
CC      -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE, THREE DISTINCT CU
CC      CENTRES KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC      3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- TISSUE SPECIFICITY: IN MYCELLA, AT A HIGHER LEVEL THAN LCCL, LCC2
CC      AND LCC3.
CC      -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOOPER OXIDASES.
CC      -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL, Z54277; CAAG1042.1; -.
DR      InterPro: IPR001117; Cu-oxidase.
DR      InterPro: IPR002355; Multicu_oxidase2.
DR      Pfam: PF00394; Cu-oxidase; 3.
DR      PROSITE, PS00079; MULTICOOPER_OXIDAS1; FALSE_NEG.
DR      PROSITE, PS00080; MULTICOOPER_OXIDAS2; FALSE_NEG.
KW      Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;
KW      Lignin degradation; Multigene family; Polymorphism.
FT      SIGNAL          1           19
FT      CHAIN           20         531
FT      DOMAIN          23         146
FT      DOMAIN          158        315
FT      DOMAIN          384        507
FT      METAL            83         83
FT      METAL            85         85
FT      METAL           128        128
FT      METAL           130        130
FT      METAL           427        427
FT      METAL           430        430
FT      METAL           432        432
FT      METAL           439        439
FT      METAL           472        472
FT      METAL           480        480
FT      METAL           481        481
FT      METAL           484        484
FT      METAL            ?          ?
FT      METAL           66          66
FT      CARBOHYD         109        109
FT      CARBOHYD         186        186
FT      CARBOHYD         231        231
FT      CARBOHYD         280        280
FT      CARBOHYD         395        395
FT      VARIANT         42          42
FT      VARIANT         119        119
FT      VARIANT         P -> S.
FT      VARIANT         H -> R.

```

FT VARIANT 246 246 R -> S.  
 FT VARIANT 256 256 P -> L.  
 FT VARIANT 261 261 P -> A.  
 SQ SEQUENCE 531 AA; 57545 MW; 737339803F75AB19 CRC64;

Query Match 6.4%; Score 198; DB 1; Length 531;  
 Best Local Similarity 23.3%; Pred. No. 1.1e-06;  
 Matches 115; Conservative 77; Mismatches 196; Indels 106; Gaps 25;

QY 78 YEMIRPFSHDIYDLEPANNVGDGSPGPTIYPRGTSVVFVNGENTSN----- 132  
 DB 25 YKEDIK--NNVADPGFORPLVSVNGVLPGLTANKG--DTLRINTNTNOLTSMRAT 80  
 QY 133 SVHLHGSF--SRAPDCAEDDTOP--GEYKDYV--PNRQARLMLYHDHMSSTAENAY 187  
 DB 81 TIRHHGLFOATLEDDGPAFYTOCPRIANLSTYEIRLHGTGTMVTHALASQYD--- 137  
 QY 188 MGQAGVYMIODP-----AEDALNLPSCYGEFDIPLVLTAKRYNADGTLFSTNG-EV 237  
 DB 138 -GLRGPVLYIDPNPHKSRVVDASVYVLMEDMYHTPAPYLRQ-----MESTNMTAL 190  
 QY 238 SSFMGDVYQVNGO-----PMPMLNV-QPKYRFRFLMAVSRFALYLAISEDET 287  
 DB 191 LSPVDSGLLNGKRGYVGGPAVPSVYVNGKRYRLRVINASVIGSTFIEGHR--- 246  
 QY 288 RLPRQVIAADGGLGPDVDTLTYSMAERWEVIDSTFAGOSIDIRNLPGADGICVER 347  
 DB 247 ---LTIVLEAD-GIRPHQPLVDSFOITVAGORISVIVEANQTA--NYWIRAPMTVAGAGINA 301  
 QY 348 EFDNTRKVRNDEVLESPD-----TSEVPAIRLDV--FPEEGNMDPANPT 393  
 DB 302 NLDPNTV---FAVLHYGAPRAEPTTEGSGAIGTAIVENHALINGAPGCG---SAPA 354  
 QY 394 DDEFTTGRANGQTINGV---TFSQYEN-----RLIRNPVROTVEIWLLEN--- 437  
 DB 355 D---VSLNLAIGRSTVGIKRFETPNNIKYEAPSLPTLKILANASNDADFTPREHTIVL 411  
 QY 438 -----NSNGWTHVHILVDPRVLSRSTARGVEPEYEAAGLKDYYWMLARREYVEA 488  
 DB 412 PHKVNIAQNHRRGADHPIHLGHVFDIV--KSLGCTPRVYVPPRDVYVKGSTGVVL--R 467  
 QY 489 HYAPFGVYMLHCH 502  
 DB 468 FKADNPGPMFVYCH 481  
 RESULT 13  
 YAK8\_SCHPO STANDARD; PRT; 622 AA.  
 ID YAK8\_SCHPO 009920;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Putative multicopper oxidase ClF7.08 precursor (EC 1.-.-.-).  
 GN SPAC1F7.08.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Murphy L., Niblett D., Harris D., Bartell B.G., Rajandream M.A.,  
 RA Walsh S.V.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: COULD BE A IRON TRANSPORT MULTICOPPER OXIDASE, WHICH IS  
 CC REQUIRED FOR FERROUS IRON HIGH AFFINITY UPTAKE. MAY BE REQUIRED TO  
 CC OXIDIZE FE(II) AND RELEASE IT FROM THE TRANSPORTER. ESSENTIAL  
 CC COMPONENT OF COPPER-DEPENDENT IRON TRANSPORT (BY SIMILARITY).  
 CC -1- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH  
 CC CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE

CC 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: STRONG, TO IRON TRANSPORT MULTICOPPER OXIDASES.  
 CC -1- SIMILARITY: BELONGS TO THE MULTICOPPER OXIDASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; 267998; CAA91955.1; -  
 DR HSSP; P37064; 1A02.  
 DR InterPro; IPR001117; Cu-oxidase.  
 DR InterPro; IPR002355; Multicopper oxidase2.  
 DR Pfam; PF00394; Cu-oxidase; 3.  
 DR PROSITE; PS00079; MULTICOPPER OXIDASE1; 2.  
 DR PROSITE; PS00080; MULTICOPPER OXIDASE2; 1.  
 KW Hypothetical protein; Glycoprotein; Transmembrane; Signal; Transport;  
 KW Oxidoreductase; Copper; Repeat; Metal-binding; Iron transport.  
 FT SIGNAL 1 22  
 FT CHAIN 23 622  
 FT DOMAIN 23 553  
 FT TRANSMEM 554 574  
 FT DOMAIN 575 622  
 FT DOMAIN 49 147  
 FT DOMAIN 194 304  
 FT DOMAIN 386 498  
 FT METAL 85 85  
 FT METAL 87 87  
 FT METAL 129 129  
 FT METAL 131 131  
 FT METAL 417 417  
 FT METAL 420 420  
 FT METAL 422 422  
 FT METAL 480 480  
 FT METAL 481 481  
 FT METAL 482 482  
 FT METAL 486 486  
 FT METAL 486 486  
 FT CARBOHYD 30 30  
 FT CARBOHYD 79 79  
 FT CARBOHYD 117 117  
 FT CARBOHYD 123 123  
 FT CARBOHYD 198 198  
 FT CARBOHYD 202 202  
 FT CARBOHYD 234 234  
 FT CARBOHYD 269 269  
 FT CARBOHYD 296 296  
 FT CARBOHYD 338 338  
 FT CARBOHYD 360 360  
 FT CARBOHYD 376 376  
 FT CARBOHYD 532 532  
 SQ SEQUENCE 622 AA; 69907 MW; A78899C7CADD2AD4 CRC64;  
 Query Match 6.3%; Score 196.5; DB 1; Length 622;  
 Best Local Similarity 21.5%; Pred. No. 1.7e-06;  
 Matches 136; Conservative 78; Mismatches 223; Indels 195; Gaps 36;  
 QY 92 DLEPAN-----MYGYDGMSPGPTIYPRGTSVVFVNGENTSNPSYHLHGSFSA 144  
 DB 36 DVDDGSGNSRWVIGVNNKRPIDPLVDVYGDQVITIKMTNSLANNRTTSLSHGLFQKFT 95  
 QY 145 FDGWAEDTQ---PGEYKDYVVPNRQARLMLYHDHMSSTAENAYMGQAGVYMIODPA 200  
 DB 96 YMDVPGSTGCEIRPGATFYVNTALQNG--TYVHSHDMQGYD---GLRTPIT--- 146  
 QY 201 EDALNLPSCYGEFDIPLVLTAKRYNADGTL---FSTNGEYSSFMGDVYQVNGQMPW 253  
 DB 147 -NALREPYDYDE-EYIISMEDWYTPPNIILVPEDEFK-----WKN--PTGAPVPDVG 195

```

OY 254 MLN-----VOPRK-YRFRFLNAVSRFALYATSESETRLPFOVIADGGLCEP 304
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 196 LFPDANATFAMPERGKTYRFRFINIG---AFNNYDWMIEDH---MTIIEVGEYTE-P 247
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 305 VDTDTLYISMAERWEVIDESTFAGSIDIR-----NLPG 339
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 248 QEVSSIHLTVARQRYSLVAKNSTDRNYAITAYMDESLEFDTIDPNVNPNTAMLSYNSDA 307
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 340 ADELGYEPEDNDKMYRFVDEVLESPDSEY-PANLRVY-----PPEEG 364
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 308 SYDLG---PDIDEIT-----SYDAELNPLYSMWTESNHSINIWPFETLGDG 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 385 GMDPNPDEDEFTFGR-----ANGQ-----WTINGVTSDEVNRLLRVPDTEIWR 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 354 ANAEIN---DSSYFPKVSIMIANSTNDGNGLEVTYGTPTNAVIFEY-GDYVDV-I 408
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 435 LENNSGWTPIVHILVDFRYSRSTARGVPEYEAAGL---KDVWL---ARREVVYE 487
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 409 IDNHDTG-KHPRHLGHTFQVLERGE-----ENAGLYSDQSHYVDNPMRDRVEIE 460
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 488 -----AHAPRPGVYMLCHNLHEDHDMMAF-----NV 517
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 461 PGSFYIRFTADNPGAMVICHIEHMEGSLATFTEAPMISISSPDEYKQCMIDGV 520
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 518 TVLGDYGVYTERID---PMEPLMRPRP---LGEFEN-----GSGDF----- 520
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 521 PTGNGNGVKNKISDLSGAPSPGEMPAWTSKAIGMACVSIACIGMSIIFYGASIH 580
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 556 ----SELAITDRIOEMASFPYQAODDAAEE 583
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 581 PVPTBELDENDLQEALENAAMELDTDKAVE 612
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
LACL_TRAVI STANDARD; PRT; 520 AA.
ID LACL_TRAVI
AC Q99044;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 01-NOV-1997 (Rel. 35, last annotation update)
DE Laccase I precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)
GN (Urishiol oxidase) (diphenol oxidase).
LN LCC1.
OS Trametes villosa (White-rot fungus).
OC Eukaryota; Fungi; Basidiomycota; Hyenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Coriolaceae; Trametes.
OX NCBI_TaxID=47662;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mycelium;
RX MEDLINE=97076915; PubMed=8975613;
RA Yaver D.S., Xu F., Goldthly E.J., Brown K.M., Brown S.H.,
RA Rey M.W., Schneider P., Walker T., Mondorf K., Dalboge H.;
RT Purification, characterization, molecular cloning, and expression of
RT two laccase genes from the white rot basidiomycete Trametes
RT villosa."
RL Appl. Environ. Microbiol. 62:834-841(1996).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) -> 4 benzenesemiquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL: L49376; AAC1686.1; -
CC InterPro: IPR001117; Cu-Oxidase.
CC InterPro: IPR002355; MultiCu-oxidase2.
CC Pfam: PF000394; Cu-oxidase 3.
CC PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
CC PROSITE: PS00080; MULTICOPPER_OXIDASE2; FALSE_NEG.
CC KX oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;
CC Lignin degradation; Multigene family.
CC FT SIGNAL 1 21 POTENTIAL.
CC FT CHAIN 1 22 520 LACCASE 1.
CC FT DOMAIN 23 148 PLASTOCYANIN-LIKE 1.
CC FT DOMAIN 160 302 PLASTOCYANIN-LIKE 3.
CC FT DOMAIN 369 491 PLASTOCYANIN-LIKE 3.
CC FT METAL 85 85 COPPER (TYPE 2) (BY SIMILARITY).
CC FT METAL 87 87 COPPER (TYPE 3) (BY SIMILARITY).
CC FT METAL 130 130 COPPER (TYPE 3) (BY SIMILARITY).
CC FT METAL 132 132 COPPER (TYPE 3) (BY SIMILARITY).
CC FT METAL 416 416 COPPER (TYPE 1) (BY SIMILARITY).
CC FT METAL 419 419 COPPER (TYPE 2) (BY SIMILARITY).
CC FT METAL 421 421 COPPER (TYPE 3) (BY SIMILARITY).
CC FT METAL 473 473 COPPER (TYPE 3) (BY SIMILARITY).
CC FT METAL 474 474 COPPER (TYPE 1) (BY SIMILARITY).
CC FT METAL 475 475 COPPER (TYPE 3) (BY SIMILARITY).
CC FT METAL 479 479 COPPER (TYPE 1) (BY SIMILARITY).
CC FT METAL 7 7 COPPER (TYPE 1) (BY SIMILARITY).
CC FT CARBOHYD 72 72 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 75 75 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 229 229 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 238 238 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 354 354 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 361 361 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 457 457 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 520 AA; 55545 MW; 046AB6D4737C60E CRC64;

```

Query Match 6.38; Score 195.5; DB 1; Length 520;  
 Best Local Similarity 22.18; Pred. No. 1.0e-06;  
 Matches 124; Conservative 74; Mismatches 203; Indels 161; Gaps 28;

```

OY 89 IYDLEPANNVYGDGMSPTIIVRGTESVYRFVNSGEN---TSPNSVHLGHSFRAP 145
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 36 VSPDGEFRQAVVYVNGRPLITGNMCDRQLVINDLHTHTWKKSSIMHGFQKG-- 93
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 146 DGAED-----TTPGEEYKDYYPNROAARMLVYHDHANSSTAENAYMGQAGVYMT 196
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 94 TNNAADGPAFINOCPISSGHSFLYDFQVPDQ--AGTEVYHSHLSTQYCD---GLRGPFV 147
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 197 QDPAEDLNLPSGGEEDILVLT-----AKRYNMDGLTFSTNGEVSSRWGD 243
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 148 YDPRDPAADI---YVDNDDTVTITLVDWYVAAKLGAPFLGADATILNKGRSPSTTTA 204
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 244 VIOVNGQPMPLWVOP--RKYRFRFLNAVSRFALYATSESETRLPFOVIADGGLLE 302
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 205 DLSV-----ISVTPKRRFRRLVSLSCDPNTF---SIDGNN---MTIIEVD-SINT 249
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 303 GYVDFTDLYISMAERWEVIDESTFAGSID---IRNLPGADGLG----- 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 250 APLVADSDIQFAQRYSFYLE---ANQAVDNYWIRANPNFGVGGINSALIRYDGA 305
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 345 --VEPEEDNDKMYRFVDEV-LESPTSVPANLRVPPPEGGMPPANPTDEFTFG 401
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 306 AAVEP--TTQTSTAPLNEVNLHPVLTAVPGS-----PVAGGVULA---IMAEFNF 354
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 402 RANGQWITGVTSDEVNRLLRVY-----PDTV---EIMRLENNNS----- 440
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 355 GTN--FRINTSTFTPIVYPLLQIISGAQNAQDLPGSYSLPSNMDIEISPPATAAP 412
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 441 GWTHPVHILVDFRYSRSTARGVPEYEAAGLKDVWL---ARREVVYEAHAPRPGVY 497

```

```

DB 413 GAPHPHLGHAFVAVY-RSAGSTVYVNDPFRDVSSTGTPAGDVVTRFR-TDNPGFW 470
OY 498 MLCCHNLJHEDHMMMAFVTVLGDYGVNTERIDMEPLMRRLPLGFEKSGDSDFSE 557
DB 471 FLHC-----HIDFLLEGFAVRED----- 491
OY 558 LATDRIQEMASFPYQAQDD 579
DB 492 -----IPDVASANPYQAQMSD 507

RESULT 15
LAC1_PLEOS
ID LAC1_PLEOS STANDARD: PRT: 529 AA.
AC 012729:
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DE Laccase 1 precursor (EC 1.10.3.2) (Benzenediol:oxygen oxidoreductase)
DE (urishiol oxidase) (Diphenol oxidase).
GN POX1.
OS Pleurotus ostreatus (Oyster mushroom) (White-rot fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Pleurotaceae; Pleurotus.
OX NCBI_TaxID=5322;
RN 11]
RP SEQUENCE FROM N.A.
RX STRAIN=FLORIDA; TISSUE=MYCELIUM;
RX MEDLINE=93514294; PubMed=7793961;
RA Giardina P., Camilo R., Martirani L., Marzullo L., Palmieri G.,
RA Sanna G.;
RT "Cloning and sequencing of a laccase gene from the lignin-degrading
RT basidiomycete Pleurotus ostreatus ";
RL Appl. Environ. Microbiol. 61:2408-2413(1995).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: 234847; CA84356.1; -
CC EMBL: 222591; CA80305.1; -
CC HSSP: P37064; 1A02.
CC InterPro: IPR001117; Cu-oxidase.
CC InterPro: IPR002355; Multicu_oxidase2.
CC Pfam: PF00394; Cu-oxidase, 3.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
KW Oxidoreductase; Signal; Copper; Metal-binding; Lignin degradation;
KW Glycoprotein; Repeat; Multigene family.
FT SIGNAL 1 23
FT CHAIN 1 529
FT DOMAIN 25 159 LACCASE 1.
FT DOMAIN 170 312 PLASTOCYANIN-LIKE 1.
FT DOMAIN 380 499 PLASTOCYANIN-LIKE 2.
FT DOMAIN 360 499 PLASTOCYANIN-LIKE 3.
FT METAL 96 96 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 98 98 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 141 141 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 143 143 COPPER (TYPE 3) (BY SIMILARITY).

```

```

FT METAL 425 425 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 428 428 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 430 430 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 481 481 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 482 482 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 483 483 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 487 487 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 57 57 COPPER (TYPE 1) (BY SIMILARITY).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 529 AA; 56580 MW; 40110AFC486882F CRC64;

```

Query Match 6.2%; Score 193; DB 1; Length 529;  
 Best Local Similarity 22.8%; Pred. No. 2.4e-06;  
 Matches 103; Conservative 58; Mismatches 180; Indels 110; Gaps 23;

```

OY 124 NSEGTSPN-----SVHLHGSFSRAPFDGNAEDTTQGE-----YKD 160
DB 71 NKGDNFQLVNQLSPTMLKTSIHWHGFFQSG--STWADGFAVYNOQCIASGNSFLYD 128
OY 161 YVYPNQARMLWYHDHAMSITAEAMVMOAGYMIQDPADALNLPBGYGEFDPLVLT 220
DB 129 FNVPDQ--AGTFWYHSHLSTGYCD---GLRGPIYVDPSPDPLSL---YVDNADPTIIF 179
OY 221 AKRY-----NADGTLFSTNGEVSFWGDDVYQVNGQP---WPMINVOPRK-YRF 264
DB 180 LEDWYHVVAQNAVLPTADSTLNGKGRFA-----GGPTSAALAVYNVESNRYRF 229
OY 265 RFLNAAVSRSFALYLATSEDSERLPFOVTAADGGLGEPVDITLYISMAEWEVYIDF 324
DB 230 RLISMCDPRFTF-----SIDGHS---LQVTEADAVNIV-PIYVDSIOIRGCGRSVLN- 280
OY 325 STFGAGSID--IRNLPGADGLGVEEPDN--TDKYMRF---VVEYLESPTSEVPANL 376
DB 281 ---ANQTVNMYWIRADPNLGGSTG---FDGINSALIRYAGATEDPPTTSSTPLEET 333
OY 377 RDVPFEGGMDPANPTDDE-----TFTGRAGQMTINVTFSVDENRLRRVNPDTV- 430
DB 334 NLVPLENPGAPGFAVPGAGADININLMAAFDVYTFELTINSFPRKAPTAPVLAQLILSGATT 393
OY 431 -----EIMRLENN-----SNGWTHPVHILVDFRLSRSTARGVEPYAAGL 472
DB 394 AASLLPSGSIYSLEANKVVEISIPALAVGGRPHFLGHRTFDYI-RSAGSTTYNPDTPAR 452
OY 473 KDVVMLARREVVYVAHY-APPQGVYMLHCH 502
DB 453 RDVVNMGTDANDVYVTRFVTDNPGPWFLLCH 483

```

Search completed: October 3, 2002, 11:02:16  
 Job time: 440 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:55:41 ; Search time 55.98 Seconds  
(without alignments)  
1801.644 Million cell updates/sec

Title: US-09-656-640A-2  
3114  
Sequence: 1 MISQAGVAGLAVIGSS.....IQEMAFNPVQADDAAE 583

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1829	58.7	602	3	Q9P8C3
2	599	19.2	475	2	Q93M03
3	501.5	16.1	637	10	Q9FTS3
4	489.5	15.7	568	10	Q23123
5	476.5	15.3	582	10	Q9AMU4
6	455.5	14.6	614	10	Q9FTS5
7	448	14.4	588	10	Q9FTS6
8	443	14.2	581	10	Q949X8
9	442	14.2	527	16	Q67206
10	431	13.8	591	10	Q9C9A4
11	402.5	12.9	536	2	Q938E6
12	401.5	12.9	502	16	Q98FM8
13	373.5	12.0	515	16	Q9CJ06
14	288.5	9.3	1662	2	P71431
15	287.5	9.2	513	16	Q9PME8
16	270.5	8.7	468	16	Q9CP61

17	259.5	8.3	463	16	Q9HXM7	Q9hxm7 pseudomonas
18	256	8.2	500	16	Q9KB49	Q9kb49 bacillus ha
19	247.5	7.9	460	2	Q9X3V2	Q9x3v2 pseudomonas
20	246	7.9	504	16	Q53858	Q53858 mycobacteri
21	245	7.9	675	2	Q9KJ08	Q9kj08 marionomas
22	237.5	7.6	561	3	Q96WN0	Q96wn0 botrytis ci
23	236.5	7.6	533	3	Q60199	Q60199 pleurotus o
24	232.5	7.5	464	16	Q9ZQV6	Q9zqv6 rhizobium m
25	231	7.4	577	10	Q9RJD5	Q9rjd5 arabidopsis
26	229	7.4	591	10	Q9AUI2	Q9aui2 pinus taeda
27	224.5	7.2	518	3	Q9HDS9	Q9hds9 polytorus c
28	223.5	7.2	609	3	Q9C1Z1	Q9c1z1 pichia past
29	222	7.1	518	3	Q9UVQ2	Q9uvq2 pycnoporus
30	221	7.1	351	2	Q93F47	Q93f47 pseudomonas
31	220	7.1	516	3	Q9Y782	Q9y782 coprinus ci
32	218.5	7.0	567	10	Q941X2	Q941x2 oryza sativ
33	218.5	7.0	589	3	Q9C497	Q9c497 glomerella
34	218.5	7.0	721	16	Q9PA43	Q9pa43 xylella fas
35	217.5	7.0	517	3	Q9UVQ5	Q9uvq5 marasmius q
36	217.5	7.0	586	10	Q9AUI6	Q9aui6 pinus taeda
37	214.5	6.9	517	3	Q12571	Q12571 basidiomyc
38	214	6.9	464	16	Q989B7	Q989b7 rhizobium l
39	213	6.8	635	3	Q96WT3	Q96wt3 candida gla
40	212.5	6.8	569	10	Q9FY79	Q9fy79 arabidopsis
41	209.5	6.7	580	10	Q9ZQW2	Q9zqw2 populus tri
42	208.5	6.7	517	3	Q9HDQ0	Q9hdq0 trimetes tr
43	207	6.6	581	3	Q96WM9	Q96wm9 botrytis ci
44	206.5	6.6	585	10	Q24044	Q24044 litrodendro
45	205.5	6.6	352	2	Q93F57	Q93f57 pseudomonas

## ALIGNMENTS

RESULT	ID	Q9P8C3	PRELIMINARY;	PRT;	602 AA.
AC	Q9P8C3				
DT	01-OCT-2000 (T-EMBLrel. 15, Created)				
DT	01-OCT-2000 (T-EMBLrel. 15, Last sequence update)				
DT	01-DEC-2001 (T-EMBLrel. 19, Last annotation update)				
DE	POLYPHENOL OXIDASE PRECURSOR.				
GN	PPOA.				
OS	Acromonium murorum.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;				
OC	Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Acromonium.				
OX	NCBI_TaxId=45278;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN-CBS 157.72;				
RX	MEDLINE=21268855; PubMed=11375170;				
RA	Gouka R.J., van der Heiden M., Swarthoff T., Verrills C.T.;				
RT	"Cloning of a phenol oxidase gene from Acromonium murorum and its				
RT	expression in Aspergillus awamori."				
RL	Appl. Environ. Microbiol. 67:2610-2616(2001).				
DR	EMBL: AJ271104; CAB75422.1; .				
KW	Signal.				
FT	SIGNAL	1	21	POTENTIAL.	
FT	CHAIN	62	602	POLYPHENOL OXIDASE.	
SO	SEQUENCE	602 AA;	66920 MW;	0303D991405228A3 CRC64;	
Query Match					
Best Local Similarity 58.7%; Score 1829; DB 3; Length 602;					
Matches 352; Conservative 83; Mismatches 131; Indels 34; Gaps 7;					
OY	5 AIAVALGLAVIGSSV-----DARSVAGRSND-----MP5GLTRPOTOLS 45				
DB	6 ALRAALALYSIKKAGQAMPKRELDIPEEPAHALAIAIVEDDPAANDLQRRSPANDLQRRSPLS 65				
OY	46 PLALYEVPLPPLPKAP-NTVPNPNTGEDILYEMETIRPFSHQIYPDLEPANWGYDGM 104				
DB	66 PAVTLQAPLSIPVKEPLFTVTNPNVNGCEIDYVEIRKHSQVFPDLGPAVLGYDGI 125				

[illegible]

RESULT	2	
Q93M03		
ID	Q93M03	PRELIMINARY; PRT; 475 AA.
AC	Q93M03	
DT	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	OXIDOREDUCTASE-LIKE PROTEIN.	
GN	AUR21.	
OS	Streptomyces aureofaciens.	
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;	
OC	Actinomycetales; Streptomycinae; Streptomyces	
OX	NCBI_TaxID=1894;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CCM3239;	
RA	Korshak J., Bistakova J., Novakova R., Homerova D., Rezuchova B.	
RT	"Cloning and characterization of a new polyketide gene cluster in	
RT	Streptomyces aureofaciens CCM3239."	
RL	Submitted (May-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AY033994; AAK61713.1.	
DQ	SEQUENCE 475 AA; 52602 MW; A2ED94AC38DD0C CRC64;	

```

Query Match          19.2%  Score 599;  DB 2;  Length 475;
Best Local Similarity 34.7%;  Pred No. 5,2e-36;
Matches 181;  Conservative 61;  Mismatches 214;  Indels 66;  Gaps 24

QY 13 LAVIGGSV----DARSVGRSTDMPSGLTKROTQLSPPLALEVPLPIPLKAPTPVN 68
    :  | | :  | | :  | | :  | | :  | | :  | | :  | | :  | | :  | | :
Db 1  MAAGAGAAALPFMDRLTGAG-SQAVASSLA---TAAAPLVPAHAMPLP-----KULKPT 52

QY 69 PNTGEDILYEMELIRPFSSQIYPDLEPANNMGVDGMSPGCTIIVPGTESVAFVNSGE- 127
    :  | | :  | | :  | | :  | | :  | | :  | | :  | | :  | | :  | | :
Db 53  SPTATSDL-YETRMBOAQVEIVKGL-MSKVRYRTDGTFFPGPTIRATGREGVVAQIMELQY 110

QY 128 NTSPTSIVHLHGSFSRAPFCGMADETTPQPELYKDYYPNCAAMLMYVHHQMSITAEWAY 187
    :  | | :  | | :  | | :  | | :  | | :  | | :  | | :  | | :  | | :
Db 111  NT---AVHLHGHAHVUSEHGLPMDITVPCGGERYRYPNNQPAASLMYVHDHAAHLEENAF 167

```

```

OY      188 MGOACVYVNIOPAEALULPGCYGEEDFPLVTAKRKNAADOTLSTNGEVSSEFGDVIQV    247
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      168 MGLHNLTYLTTHNETTLPDS--GPYDVLPIRRAAREADOSTLLTYTRPSCPH---MLV    221
OY      248 NGQPMPMLNVOPKRTFFFLAANVSRSALEYLATSESETRLRPQVYAAGGLEGVDT    307
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      222 NGRKPDPVOVAKARFRFYVNACNRRYVKLFACG-----IEFTQIGTGGLFLEPVQO    275/
OY      308 DLYLSMAERMEVVIDEFTF-AGOSIDIRNLPGADGLVEEPENDTKVRFEVDEVLES    366
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      276 SELMLLGERADIYVDFSRYKVGDOSVLEN-PGAOSI-ERPE-----VMRF-DIVRTA    325/
OY      367 PDTSEVPANLEDRPERPEGGMNDPANPPDDDEFTTGRANGOWITINGCTFFSDVENELLANVP    426
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      326 PUYSVPRRLTTYR-PQ-----PPPYVERDEFI-RTPPMATINGOSID--PNWVDITAK    375/
OY      427 RDIYVEIMRLEN-----NSNGWTIPRHILHVDVFRSLRS-TARGVEPEAAGLKDV    475
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      376 LGTEVETVTVRNVEAPAAEGKKDFHLHM-HSFPTHLYTFVEFLERNCKPAGTRD--LGDKT    431/
OY      476 WMLAREVYVVEAHYAPRPGGYUMLHCCHLHJEDHDMMAAFNVP    517
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      432 VTLGGPDVTKIAMTWGPTGTGQYLTHCHODLGSSSGGOMCRIDI    473

```

SEQUENCE	637 AA:	70919 MW:	692D39F10FD26A96 CRC64:
SO	SEQUENCE	637 AA:	70919 MW: 692D39F10FD26A96 CRC64:
KM	coat protein.		
DR	PROSITE: PS00080; MULTICOOPER_OXIDASE2; UNKNOWN1.		
DR	InterPro: IPR002355; Multicox_oxidase2.		
RL	EMBL: AP002860; BAB18287.1; "		
RT	Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.		
RT	clone:00409B08.*;		
RA	"Oryza sativa nipponbare(Ga3) genomic DNA, chromosome 1, PAC		
RC	Sasaki T., Matsumoto T., Yamamoto K.;		
RP	STRAIN-CV. NIPPONBARE;		
OC	SEQUENCE FROM N.A.		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoideae; Oryzaceae; Oryza.		
OX	NCBI_TaxID=4530;		
RN	[1]		
OS	Oryza sativa (Rice)		
DE	P0409B08.14.		
DE	PUTATIVE SPORE COAT PROTEIN-LIKE PROTEIN.		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
AC	O9FTS3.	PRELIMINARY;	PRT: 637 AA.
ID	O9FTS3		
RESULT	3		

	Query Match	Score 501.5;	D8 10;	Length 637;
	Best Local Similarity	28.7%;	Pred. No. 1.3e+28;	
Matches	177; Conservative	80; Mismatches	194; Indels	165; Gaps
OY	20 SVDARSVACRSTDMPSGLTKRTQTQSLPPLALXYEVLPIP-----PLKAPVTNPENPNIG	72		
	:::::	:	: :	
Db	57 SYVSSSRMCWMAADDPN---DDEXTTP----DHLPAPAAGRMRPVN--TSLNLTRY	104		
OY	73 EDLYKEMEIRPF-----SHOQLPLEBANMYGYGMS-----PG	107		
	:::::	:	: :	
Db	105 VDSLPIAKIKRGYGIRHGRVPFIKLITGMYSKTWFHRMPPTPRVVY-GQSLSGTATFFGC	163		
OY	108 PTIIIVPGTESVVRVN-----SGENTSPNSVHLHGFSRAPFDGA	149		
	:::::	:	: :	
Db	164 PTVARQGVAFLAVEMQNHLIPDAHIILPWDEKVPTAIPKKCGGVTVVHALHGCAHPREEDCHA	223		
OY	150 -----EDTPOBE---YKRKYYPNRQAAMYLWHDAQSITEMANMGAGAYVMIDPP-A-200			
	:::::	:	: :	
Db	224 FAMETRDFEAENSJTWTRKTYTTPNOAQEGMLMTHDALCLTRYSLAGLAAYVEIKPEL	283		



```

QY 201 EDALMPCGYCEEPDIPVLTAKRKNADGOTLESTGEVS-----SPWGDVIOVNGQMP 253
Db 284 EDPMNLPC--GDHDLHLVIADRE-----FTYNGSISIDREMKPEYFGVLTVNGKAMP 334
QY 254 MLNVOPRKRYRFRFLNAVSRFSFALYLTSEDSETRLPFOVIAADGGLLEBPVDTLYIS 313
Db 335 YLSVORRRYRRLRINASNARFENVTL-----SNGALPFTYIGSDSSVLSRPVIVSNLVLS 369
QY 314 MAERKEVYIDSTFAGOSIDIRNL-----FGADSLGV-EPEEPDNTDKVMFVY---DEVL 364
Db 390 PAELFDVYIDVDSRLPAAMTELEMTANTAPRPPNPNPTDNLGG--KVMLEKYAGKQVD 447
QY 365 ESPPTSEVPALIRLVPPEEGGNMPPADPT-----DDFTFGGANGOWINVTSFQVE 418
Db 448 DMPKSKVPEH--GVPAASVALRPPPTTKRIVLYENOT-----APGNLYINGRLRLEDP 499
QY 419 NLLRLNVPRD-TVEIWRLENNSGWTHPVHILHLYDFVRLSRSTARGVEPEAAGLD-- 474
Db 500 --VTEPDKSTTELMQV-INLTDGNHPLHLIATFOAIKMTKEGQVPEKDKMIKNNT 555
QY 475 -----VVMILAREVVY-----VEALYA-PF-----PGV 496
Db 556 ATCNLDHANGVPPVPPEEKTKMANAVKIPPEMTSVVAFLRLEANQYPFDATTEPG- 614
QY 497 YMLCHNLHEDDHMM 512
Db 615 FVYCHLHEDHDNAMI 630

RESULT 4
023123 PRELIMINARY; PRT; 568 AA.
AC 023123:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F19G10.5 PROTEIN.
GN F19G10.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_taxonomy:3702;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV. COLUMBIA:
RC Federapfel N.A., Palm C.J., Conway A.B., Kutz D.B., Conway A.R.,
RA Au W., Ataulo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
RA O'J O., Osborne B.L., Shum P., Sun H., Toriumi M., Vytotskaia V.,
RA Yt G., Ecker J., Theologos A., Davis R.W.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF000657; AAB72167.1; -;
DR InterPro: IPR001117; Cu-oxidase.
DR Pfam: PF00394; Cu-oxidase.1.
SQ SEQUENCE 568 AA; 64355 MW; C4314C889576E35E CRC64;

```

Query Match	15.7%	Score 489.5	DB 10	Length 568
Best Local Similarity	28.3%	Pred. No. 8.3e-28		
Matches 154; Conservative	77;	Mismatches 181;	Indels 133;	Gaps 24.
OY	82	IRPFHQI-----YPDEPAMWGDGMS-----PGGTIVPRGESVFRFNSG	126	
Dd	50	IKPASLQIGMFSTKMKFHRDL-PAPRFVAYGRSRSKATVPGLTILEVYGVDIDVYWRNLL	108	
OY	127	ENTS-----PNSVHLHGSTRAPFDGMAE-----DTDQGEYKD	160	
Dd	109	PKSHILPMDPTISPATPKHGGIPRTVVHLHGCIHEPTSGNAAMAFAGRETPRKWTKT	168	
OY	161	YYPRQAARMYLWDHAMSIFAENAYNGAQAVNYMIDPA-DDALNLPGCYEEDFILPLV	219	
Dd	169	LATEKKOOPGNMMYIDHAAIGLTRVNLGLGVATILLRHVAVSPOLPTPG-DDEFDPILT	227	

```

QY 220 TAKRNMNDGLF-----STNGVSSFW-----GDVIOYNGQBPWMLYANQPKRYFRFLMAA 270
Db 228 FDRSRFRKSGYSTYMATAGNNPDSIHPOWODEYFGVDIYVNGAMPMLNRRKRYRRRLTMA 287
QY 271 VSRSEATLATSEDESETRLPQVLAADGGLGEGVDVDITYLSAAEWEVVIDFSTFAGQ 330
Db 288 NARFFKFFFSNG-----LDFIVGSDSAYLSKRYVMKSTILBSSELYVDVYVDFYKPSR 344
QY 331 SIDRN--LPGADGLGVEPEPDNTDKVMRFVDEVLSEPDTSVEPANLKDVPPEECN- 386
Db 342 TVVLANADAPYPPSG--DPVNEENGKVMKFIINSESD-DTCTIPKLLINYPADVSNA 397
QY 387 -----WDPANTDDETFFFGRANQOWTINGTESDVENRLLRANPRTDVEIWLRENN 438
Db 398 VLTRYISMTEVYSNSDEPHTLL-----VNGLPY--EAPVTEPKSGTTEWEVINL 446
QY 439 SNGWTHPVHILVDFRVLRSST--ARGVEPE-----AAGLKDVV-- 476
Db 447 TED--NHRPLHILGLFKVEQYFALLAAGLEERKECMTQONDAVKKQISKYARCKKTAVTAH 505
QY 477 ---W-----LARREVVVEAHYAP--PGVYMLHCHNLIHEDHDMMAA 514
Db 506 ERGKNVFKMAMPGHVTRILVRFYSYHTNASY-PEDPTQPEG-YVYHCHILIHEDNNMAMP 563
QY 515 FNVTY 519
Db 564 LKVIT 568

```

Query Match	15.3%;	Score 476.5;	DB 10;	Length 582;
Best Local Similarity	28.9%;	Pred. No. 7.8e-27;		
Matches 156;	Conservative 60;	Mismatches 165;	Indels 159;	Gaps 24;
0y	88	Q1YPLEPANNWGVIGDMS-----PGPTIYVRGIESVYRFVN-----	124	
DB	80	KFHRL-PASVITFVGTSAAATPFGPTIEAGVPLSVTWQNYLDPARILPMDPTVPA	138	

QY	125	-SGENTSPNSVHLHGSGFSRAPEDD-	MAEDTTPQGYKDYVPPNRQ	167
Db	139	IPRRGCVPTVYHLHGCAHPPOSDSASAPWTFAGGEGFPANSTPT	YTPNMQ	190
QY	168	AARMLMYHDHAMSITTAENAVYMGQAGVYMIODPAEDA-LNLPSCYGEEDIPVLITAKRYNA		226
Db	191	SPGLWYHDHMLGLTRANLLAGLLGAVYIRNPVAEPLGLPCG-DEFDRLMLADRSFYA		249
QY	227	DGTLFSTNGEV-----SFFMGDIYQVNGQWPMNLQPKRYRRRPLMAAVSRFAL		277
Db	250	DGSIYMTNTGILPINHPOMQDEYGEALTYNGKAMPPLAARRRYRFRRIINTSNARENL		309
QY	278	YLATSEDESETPLPROVIAADGGLLEGAPVDITLITSAERWEYIDPSTFAGOSIDIRN-		336
Db	310	SLTNG-----LPLTVVGSDTNLYLSKPTVLAASLLVSAEFDVVDVDSQSTSPAELVNT		363
QY	337	--LPGADGLGYPEPEDNTDKMYRKYVDEVLESPTSEVPANLRD-----VP		360
Db	364	APYPYPDG---QAQPDNLGKMYKVFISPA-KAKDTSRYPAKLIDYAVAAEEAVQRYIV		419
QY	381	FPEGGMNPMPANPTDEPTFCFRANGOWTINCVPFSVDENRLLRVPR-DTVEIWRLENN		439
Db	420	MYEYDATTGNT-----HLXNKRLEDPAT-----ETPRGTTEWEVYINT		463
QY	440	NGWTHPRVHILVDFRVLSRSTARGVEPEYA---AGIKDVY-----		476
Db	464	PD-NHPLHLHLATFQA---TVRGLVDDEDAFKGMAKLINDAVRCNVSRAHAGEVAVPEH		519
QY	477	---MLAREVYVYVEAHYA-----PF-----PGVYMLHCHNLIHEDHDM		512
Db	520	EKGW---KNVVKIAPGVYTTIYVFFWYDSSKPIPFDAATAEPG-YVICHILHEDNAMT		575
RESULT	6			
Q9FTS5		PRELIMINARY:	PRT:	614 AA.
AC	Q9FTS5			
DT	01-MAR-2001 (TREMblrel, 16, Created)			
DT	01-MAR-2001 (TREMblrel, 16, Last sequence update)			
DT	01-JUN-2001 (TREMblrel, 17, Last annotation update)			
DE	PUTATIVE SPORE COAT PROTEIN-LIKE PROTEIN.			
GN	P0409B08.12.			
OS	Oryza sativa (Rice).			
OC	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehharotidaeae; Oryzaeae; Oryza.			
OX	NCBI_TaxId=4530;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV, NIPPONBARE;			
RA	Sasaki T., Matsumoto T., Yamamoto K.;			
RT	"Oryza sativa nipponbare (CA3) genomic DNA, chromosome 1, PAC			
RT	clone:P0409B08."			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AP002860; BAB18285.1;			
DR	InterPro; IPR002355; MultiCo, oxidised.			
DR	PROSITE; PS00080; MULTICOOPER_OXIDASE2; UNKNOWN_1.			
FM	Coat protein.			
Q0	SEQUENCE 614 AA; 68321 MW; 800CE68563E305193 CRC64;			

```

Query Match      14.6%  Score 455.5; DB 10; Length 614;
Best Local Similarity 27.6%  Pred. No. 3e-25;
Matches 171; Conservative 75; Indels 155; Gaps 30.

QY 20 SYDARSVAGSRDMSGLT-----KQQTQLSPPLALKEVPLRPPLKAPTV----- 66
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 18 SYISKSRMGWASDDDDDDYTPAADPSPTAPARPLURSTADLEKRYVDPQMARIGYGINDFG 77
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY 67 -PNPTGDEDILYEMEIREFSHQIIPDLEPAMWVGIDGMS-----PGPTIIVPRTSEYV 119
    |||||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 78 FVPV-TNLTIGMYNK-----TWQFRHDDMPPTVEYV-GGSLQATATFPGPTIIVARNVPLY 130
    |||||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

```

0Y 120 VRFVNSGENTS-----PVSYLHOSFSRNPEDGWA-----EDTQPG 156
Db 131 VTWENHLEDAIILEWDPVPVPAIILKNGCVPLVYHLHQAQAQPSDGAIFMFRDFAENG 190
QY 157 E---YKDYUYNRK--AARMLYHDHAMSITAEANVMQAGVYMIQD-ABDALNLPSCYG 211
Db 191 STWQKTYTYVNOVPAAGINIMYHHAALGLTRASLLAGLAAIYIEMPELMFPLPS--G 248
QY 212 EFDPLVLTAKRYNADGLF-STNGEYSS-----FMGDVIOVNOQPMPLVNOQRKY 262
Db 249 EFDLHLVADKREKFNVDGITEFMDTIGAVPSVHPQOPREYFGEVITIVNSKAMPFOAVQORRY 308
QY 263 RFRFLNAAVSRSPALYATSEDSSETRLPFOVIAADGGLGEPVDFTDLYISMAERWEV1 322
Db 309 RLRLINASNARYLNIREFSNG-----LPFVYIADATVLTSRPVYSNLLSPAEIPEYIV 362
QY 323 DESTFAGS---IDIRN--LPGADGLGVEEPEPNTOKVYRFV-----DEVLESPD 368
Db 363 DFLSVLVNPNATDIELNSAPYPPPTGPANATLDG--KVMAFNVSAAKQYGDDBKMOPE 420
QY 369 TSEVPANLRDVPPEGGNMDDA-----NPTDDETFEFGRANQMTINSGVTSDE 418
Db 421 NSTVPEI-GVDFPAKVLTALPTMKTRYIVLEYENNTSMDPNTAKTM--LYINGRLBEDP 477
QY 419 NRLLRNV-RQTVELWRLNNNSQWTRPHVHLVDFKLSSTARKVEPEAAGLK---- 473
Db 478 T---ETPISGTELMHWINLTDP-NHPLHNLHLEFQAVOMLOLVDPDTERKSCMLKHNDT 522
QY 474 -----DVYMLAREVVYVEAHYA-----PF----- 493
Db 533 FACNLDDQAVGALQVPVEEEXTM---KNVYKIPRAYTYSVVAFRLVHNMMNPRFDATAA 569
QY 494 PGVYMLHCHNLIHEDHMM 512
Db 590 PG-YVYHCHILHDHEDNMI 607

```

RESULT	7		
Q9FTS6			
ID	Q9FTS6	PELLEMINARY;	PRT; 588 AA.
AC	Q9FTS6;		
DT	01-MAR-2001 (TREMBLrel, 16, Created)		
DT	01-MAR-2001 (TREMBLrel, 16, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel, 17, Last annotation update)		
DE	PURATIVE SPORE COAT PROTEIN-LIKE PROTEIN.		
CN	P0409B08.11 OR P0044F08.29.		
OS	Oryza sativa (Rice).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tr		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae		
OC	Ehretiaceae; Oryzaceae; Oryza.		
OX	NCBI_TaxId=4530;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV, NIPPONBARE;		
RA	Sasaki T., Matsumoto T., Yamamoto K.;		
RT	"Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1,		
RT	clone:P0409B08.";		
RL	Submitted (Aug-2000) to the EMBL/GenBank/DBD databases		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV, NIPPONBARE;		
RA	Sasaki T., Matsumoto T., Yamamoto K.;		
RT	"Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1,		
RT	clone:P0044F08.";		
RL	Submitted (Nov-2000) to the EMBL/GenBank/DBD databases		
DR	EMBL; AP002860; BAB18284.1; -		
DR	EMBL; AP002903; BAB21194.1; -		
DR	InterPro: IPR002355; MultiCu_oxidase2.		
DR	PROSITE: PS00080; MULTICOPPER_OXIDASE2; UNKNOWN_1.		
CO	Coat protein.		
SO	SEQUENCE 588 AA; 64290 MW; 6C6CB995060C4A82 CRC64;		





DR EMBL: AY053392; AAL15149.1; -  
SQ SEQUENCE 536 AA; 58532 MW; DAE43663DIB12B74 CRC64;

Query Match 12.9%; Score 402.5; DB 2; Length 536;  
Best Local Similarity 28.4%; Pred. No. 2e-21;  
Matches 140; Conservative 55; Mismatches 199; Indels 99; Gaps 18;

QY 100 GYDGMSPGPTIIVPRGTEGVRFVNS-GENTSPNSVHLHGSFGRAPDPAEDTTPQGEY 158  
DB 68 GYGNGLGFAVOLHRSKSTVVDHNOAEDT---TIHMGLEITPGYDGGPQIIPAGCT 124  
QY 159 KDYYNROAARMLWYHDHAMSITTAENAYGQAGVYMIODPAEDALNPSGYEPIPLV 218  
DB 125 RYVTFPEQRAATCMIHPRHCKTGQVAMGLAGLVIEDDEIRKLRLPKQMGIDDPVY 184  
QY 219 LIAKRYNAGTGL---ESTNGEVSSFMGDIYQVNGQPMPLNVQPKRYRFRFLAANVSRF 275  
DB 185 IQDKRFSADGQIDYQDIDMTAAVGMFGDILLTGATYPOHSAPKGLRLRLNGCNARS- 243  
QY 276 ALYLTASEDESEETLPROVIAADGGLLEGPDVDTLYISMAERWEVVIDSTFAGOSIDIR 335  
DB 244 -LNIASDR-----PLXVIASDGGGLAEPKYVELPLMGEREVLVDISD--GKAFLV 296  
QY 336 NLPGAD-GLGVEPEFDNTPKVMRFVDEVLES--PDT-SEVPANLRDVPFPEG- 384  
DB 297 TLIPVSGMGMAIAP-FDKPRHVMRIQPAITASGLTLPDTFTTMA---LPSEGLVRL 351  
QY 385 -----GNMPPA----- 390  
DB 352 KLSMDRLDMGMOMLKKYGAQAMSGMDHDSNNAHMGNGMGMDHGMDSGMNHG 411  
QY 391 ---NPTDDETFEGRANGOMTIVTFSDVENRLRNVPDVEIRLENNNGMHPVH 447  
DB 412 ANGNMNHGKRFDPHNN-----FINGVFD--MKNRPMRAQGRHERRVVISGVGMHLRPH 466  
QY 448 IHLVDFRVLRSRSTARGVER-YEARGLKDYVWL--ARREVVYEAHAPPGVYMLCHNL 504  
DB 467 IHGTQRIISEN---GKAPAAHRTGKMDYRVGEGISEVLVKFDHDAKPEHAYMAHCHLL 523  
QY 505 IHEDHDMMAAFNV 517  
DB 524 EHEDTGMLGFTV 536

RESULT 12  
Q98FW8 PRELIMINARY; PRT; 502 AA.  
AC Q98FW8: 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, last sequence update)  
DE 01-OCT-2001 (TREMBlrel. 18, last annotation update)  
GN ML13586.  
OS Rhizobium loti (mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Mesorhizobium loti."  
RL DNA Res. 7:331-338(2000).  
DR EMBL: AP003002; BAB50448.1; -  
DR InterPro: IPR001117; Cu-oxidase.  
DR Pfam: PF00394; Cu-oxidase; 3.

KW Complete proteome.  
SQ SEQUENCE 502 AA; 53995 MW; AE96D894241B36CE CRC64;

Query Match 12.9%; Score 401.5; DB 16; Length 502;  
Best Local Similarity 29.1%; Pred. No. 2.2e-21;  
Matches 162; Conservative 57; Mismatches 232; Indels 105; Gaps 25;

QY 7 GAVALGIATVIGSSVSDARSAG-RSTDMPS-----GLTKQTQTSPP-LALXYEY 53  
DB 5 GFLAAGLAAY---LVSERAVAQMKMDMPGMDGAGHDMGAAPQGVVPPGEVYREL 61  
QY 54 PLRIPPLKAPNTVPNNTGEDIITYEMEIRPFESHOIYPLDELPANMYGYDMSGPTIIVP 113  
DB 62 PL-----LANEASRGLEFATILTAGSATARFAKGL-----DPIILAYNTSGPLIEAV 110  
QY 114 RGTSEVYRVNSENTPSPNSVHLHGSFGRAPDPAEDTTPQGEYKDYYPNRQA-ARML 172  
DB 111 EGRVREITTFANRIANRA-STIHHGMPVPADODGNPMDPVAATGTDRTYSFDLEASAGSY 169  
QY 173 WYHDHAMSTAEANAYGQAGVYMIODPAEDALNPSGYGEFDIPLVITAKRYNADGTLFS 232  
DB 170 WYHPRHCKTAEOVYRQLAGAFVYKPRADP--TPAYG--DTVLVPTDLRLADGTLPP 224  
QY 233 T-----NCEVSSFMGDIYQVNGQPMPLNVQ--PRKYRFRFLAANVSRFALYATSEDS 285  
DB 225 NMTEDLMNGV-----GDHVLVNGQKNPTLVPFGAKRRFRFYATNAR-----FLRLSFDG 276  
QY 286 EYRLPROVIAADGGLLEGPDVDTLYISMAERWEVVIDSTFAGOSIDIRNLPGADGLV 345  
DB 277 AS---MTIIGTDGGLLEAPVAAGDILSPARLELVVSF-----DKPGAAL-T 321  
QY 346 EPEFDN-----TDKVMRFVDEVLESPTSEVPANLRDVPFPEGGWDP 389  
DB 322 TLIDYDGMGCRPADAGTLTLTVNSQTPADVPLPLDRLPAOL----- 368  
QY 390 ANPTDDETFEG-----RANGQ--MTINGVTFSDVENRLRNVPDVEIRLENNNSNG 441  
DB 369 GAVAVSRFRFETFTMAAMNSGMEMGLINGAFL-DWQ-RIDVAKKGOVELMIVHEAD- 425  
QY 442 WYHPRHCKTAEOVYRQLAGAFVYKPRADP--TPAYG--DTVLVPTDLRLADGTLPP 501  
DB 426 MDPRFVHGTQFOVVEHERGNISKPAYRAWKDTVVAVGEAVRLILR-QDRPGPMYHC 484  
QY 502 HNLHEDHDMMAAFNV 517  
DB 485 HILHEQLGMMGVYDV 500

RESULT 13  
Q9CJ06 PRELIMINARY; PRT; 515 AA.  
AC Q9CJ06: 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
GN PM1938.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM70;  
MEDLINE=21145866; PubMed=11248100;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida pm70."  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
DR EMBL: AE006230; AAK04022.1; -  
DR InterPro: IPR001117; Cu-oxidase.  
DR InterPro: IPR002355; MulticCu\_oxidase2.  
DR Pfam: PF00394; Cu-oxidase; 2.

DR PROSITE: PS00080; MULTICOOPER.OXIDASE2; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 515 AA; 58572 MW; 1665D9B6C7C6743F CRC64;

Query Match 12.0%; Score 373.5; DB 16; Length 515;  
 Best Local Similarity 26.4%; Pred. No. 2,6e-19;  
 Matches 142; Conservative 69; Mismatches 200; Indels 127; Gaps 22;

```

OY 33 MPGLKRTQTSPLALVEPLPIPLKAPNTVPNPTNGEDILYYEMELRPSHOIYPP 92
DB 51 MPGLS--LNGILPKLA-----NOSTQAGLFTYTLKAEPIKIRI-AD 89
OY 93 LEPANVGVGDMSPGPTTIYPRGTSVVRVNSGENTSPNVHLHGSFRAPEGMAEDT 152
DB 90 NKTEERWANGQLPQPIQIEFGEDTVEIERIN--HLRQSTVIMHGLDVNEADGNPMDM 147
OY 153 TQGEYKDYIYPRNQ--AARMLYHDAHMSITAEWAYGQAGVYIIDDPAEDALNDSGYG 211
DB 148 VEPQKQVYRFTLPQSGAGTYWYHPHHDVSEQVYKGLAGTFVYAKNDPLAHLPEOH- 206
OY 212 EEPDIPVLAKRYNADGTLST-----NGEVSSFWGVDYIQVNGQPMPLNQPRKYRFR 265
DB 207 ----WISDLRLNADGTLIPANTMLDMNGRE---GEFVLINGQYQPOIQTNE-RIR 256
OY 266 FLNAVSRSEFALYATSESETRLPFOVLAADGLEGP-VDTDTLYISAEEMEVYIDF 324
DB 257 LMMATGARFRRLNI-----PEVKMIYVGTBEGGLEKPRAPDELLITPAEVEVIM-- 367
OY 325 STFGAGSIDIRNLPGADGLGVEPEFNDTKVMRFVYDEVLESP-----DTSEV 372
DB 308 ---VGETQGVNL-----QSGYVDRKRM-----VOEQPKDLILATIQKSEPVQI 350
OY 373 PANTRVVPEEGNMD-----PANTDDEFTFG 402
DB 351 PESTRLP-----NMDEPKVOIRSEKMMNTNMPNMGHGHATTPTDNPPIPM-- 403
OY 403 ANQWITNGVTSDEVNRLRNVRDVTVEIWRLENNNGVTHILVDFVLSRSTAR 462
DB 404 MNGMFLINGTFD--KNRIDFVAKLNEVEQWEIFNESH--MDHFFHLGTFEVIQHTLNG 460
OY 463 GVEPYPAAGLKDVMYLAARVVYV---EAHYAPFPGVYMHCHNLHEDHDMAAEYV 517
DB 461 KTFEPGKRALKDVMYLAARVVYV---TGKMYHCHILHEMLGMGMKEVY 514

RESULT 14
P71431 PRELIMINARY: PRT; 1662 AA.
AC P71431;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MOFA PROTEIN PRECURSOR.
GN MOFA.
OS Lepidoptrix discophora.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Lepidoptrix.
OX NCBI_TaxID=89;
RN [1]
RN SEQUENCE OF 1-1150 FROM N.A.
RA STRAIN-SS-1;
RA Corstjens P.L.;
RT Identification and Molecular Analysis of the Lepidoptrix discophora
RT SS-1 mofa Gene, a Gene Putatively Encoding a Manganese Oxidizing
RT Protein with Copper Domains";
RN Geomicrobiol. J. 14:91-108(1997).
RN [3]

```

RP SEQUENCE FROM N.A.  
 RC STRAIN-SS-1;  
 RA Corstjens P.L.;  
 RL Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: Z25774; CAAB1037.1;--  
 DR InterPro: IPR001865; RIBosomal\_S2.  
 DR PROSITE: PS00962; RIBOSOMAL\_S2\_1; UNKNOWN\_1.  
 KW Signal.  
 FT SIGNAL. 1 33 POTENTIAL.  
 FT CHAIN 34 1662 POTENTIAL.  
 SQ SEQUENCE 1662 AA; 174293 MW; 639EE236600D9246 CRC64;

Query Match 9.3%; Score 288.5; DB 2; Length 1662;  
 Best Local Similarity 19.4%; Pred. No. 2,8e-12;  
 Matches 133; Conservative 64; Mismatches 144; Indels 345; Gaps 20;

```

OY 19 SSVDANSVAGRSTDMPSGLTKROTQSPPLALVEPLPIP-----PLKAP 63
DB 75 SEADALFAGKFTG-----RSATDQKALRKFIIDPLPGKLSLADGVQRTIVAVP 128
OY 64 NTVPNPTG-----EDILYYEMELRPSHOIYPPLE-PANVGV-----DG----- 103
DB 129 AKWINDQTGAATSD--YFELALVEYKQKLHSDLKMPYTLRGVQLSDVDTGLYLYPNS 186
OY 104 -----MSP-----GPTIYPRGTSVY 120
DB 187 SKGAFANADADADAKPVYINAVDAKGFTGAKVQARVYMERHMGPIIQAQKGPTRL 246
OY 121 RYVN-----SGENTSP-----NSVHL 136
DB 247 KFNLLPGRAETTVGADGVQYATARNGDIPLDLKSIHAGLPGDFTFQTNISHL 306
OY 137 HG-----SESRAPFGMAEDTTPG 156
DB 307 HGGDTPIISDGTPTPHOWITPLEEANAANKALVNOGIDPEFLPSFLGASQANVPDMPG 366
OY 157 E-YKDYYPNRQAARMLYHDAHMSITAEWAYGQAGVYIIDDPAED-----ALNL 206
DB 367 AGASTYTFPQGQARMLYHDAHMSITAEWAYGQAGVYIIDDPAED-----ALNL 206
OY 207 PSCYGEFDIPVLTAKRY-----NADGTLFSTGEVSSF----- 240
DB 427 VLPPADDTIPLVLTDFVPADVALQDARWNTSAMGSESDSWPPHYETVQDPNONGFN 486
OY 241 -----MGDVIQVNGQPMPLNQPRKYRFR 265
DB 487 SVGRMHMGPFVFPFAMVDLPSEYGDVITVPEAMMDPLVNGVAYPTIELDPKYRMR 546
OY 266 FLNAVSRSEFALYATSESE-----TRL----- 289
DB 547 VLNASNDREFNISLFLVADEAQRNLDPILGATEVKWDAVASATPCAGVATRAVATDGS 606
OY 290 -----PFOVLAADGGL-----EGPVD--- 306
DB 607 YCTPEWPTDNRPGVSPPAQSPFEOIANEGGLLPKVAELIAPTPVYQLDGRITVLN 666
OY 307 --TDLYISAEEMEVYIDSTFGAGSIDIRNLPGADGLGVEPEFD----- 350
DB 667 VLTGTLGNARADVLVDLSAYAGKTLIYVNDSGAPVPAGDPRNDYFTAVGDSQDAGCA 726
OY 351 -----NTDKVMRFVYDEVLESP 367
DB 727 EDTKPGVGPNTRTMQLKVAALITTP 752

RESULT 15
O9PME8 PRELIMINARY: PRT; 513 AA.
ID O9PME8
AC O9PME8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

```

DE PUTATIVE PERIPLASMIC OXIDOREDUCTASE.  
GN CJI516.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
OC Campylobacter.  
OX NCBI\_TaxID=197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCTC 11168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Peltwell T., Holtroyd S.,  
RA Juels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajadwaram M.A., Rutherford K.M., Van Vliet A.H.M.,  
RA Whitehead S., Barrall B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
RT reveals hypervariable sequences.";  
RL Nature 403:665-668(2000).  
RL EMBL: AL139078; CAB73936.1. -  
DR InterPro: IPR001117; Cu-oxidase.  
DR InterPro: IPR002355; Multicou\_oxidase2.  
DR Pfam: PF00394; Cu-oxidase. 1.  
DR PROSITE: PS00080; MULTICOPPER\_OXIDASE2; 1.  
KW Complete proteome.  
SQ SEQUENCE 513 AA; 59076 MW; 019CBAE21B7A1555 CRC64;

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:16 ; Search time 25.3 Seconds  
(without alignments)  
562.851 Million cell updates/sec

Title: US-09-656-640A-2

Perfect score: 3114  
Sequence: 1 MISQATGAVALGLAVIGSS.....IQEMASFPYQADDDAAEE 583

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PC10S.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfill1es1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3114	100.0	583	4	US-09-401-476-2
2	2540	64.5	572	4	US-09-401-476-4
3	254.5	8.2	616	1	US-08-749-882A-2
4	254.5	8.2	616	2	US-08-539-134-2
5	253.5	8.1	616	2	US-08-991-531-2
6	253.5	8.1	616	2	US-09-032-315-9
7	253.5	8.1	616	2	US-08-993-318A-9
8	253.5	8.1	616	3	US-09-028-887-2
9	253.5	8.1	616	4	US-09-399-886-9
10	253.5	8.1	616	4	US-09-396-260-9
11	253.5	8.1	616	4	US-09-518-901-2
12	253.5	8.1	616	4	US-09-576-281-9
13	253	8.1	620	1	US-08-940-661A-2
14	253	8.1	620	2	US-09-083-485-2
15	253	8.1	620	2	US-08-939-218A-2
16	252.5	8.1	616	5	PC1-US95-06816-2
17	241	7.7	573	2	US-08-991-531-1
18	241	7.7	573	2	US-09-032-315-10
19	241	7.7	573	3	US-08-993-318A-10
20	241	7.7	573	3	US-09-028-887-1
21	241	7.7	573	4	US-09-399-886-10
22	241	7.7	573	4	US-09-396-260-10
23	241	7.7	573	4	US-09-518-901-1
24	241	7.7	573	4	US-09-576-281-10
25	241	7.7	620	1	US-08-706-037-27
26	241	7.7	620	1	US-09-005-397-27
27	241	7.7	620	5	PC1-US95-06815-2

28	220	7.1	516	3	US-08-689-421-29	Sequence 29, Appl
29	220	7.1	516	4	US-09-389-528-29	Sequence 29, Appl
30	220	7.1	516	4	US-09-181-827A-29	Sequence 29, Appl
31	213	6.8	529	1	US-08-172-318A-14	Sequence 14, Appl
32	213	6.8	529	2	US-08-706-037-25	Sequence 25, Appl
33	213	6.8	529	2	US-09-005-397-25	Sequence 25, Appl
34	213	6.8	529	2	US-09-032-315-5	Sequence 5, Appl
35	213	6.8	529	2	US-08-993-318A-5	Sequence 5, Appl
36	213	6.8	529	4	US-09-399-886-5	Sequence 5, Appl
37	213	6.8	529	4	US-09-396-260-5	Sequence 5, Appl
38	213	6.8	529	4	US-09-576-281-5	Sequence 5, Appl
39	202.5	6.5	539	2	US-09-032-315-1	Sequence 1, Appl
40	202.5	6.5	539	2	US-08-993-318A-1	Sequence 1, Appl
41	202.5	6.5	539	3	US-08-689-421-27	Sequence 27, Appl
42	202.5	6.5	539	4	US-09-399-886-1	Sequence 1, Appl
43	202.5	6.5	539	4	US-09-396-260-1	Sequence 1, Appl
44	202.5	6.5	539	4	US-09-389-528-27	Sequence 27, Appl
45	202.5	6.5	539	4	US-09-181-827A-27	Sequence 27, Appl

## ALIGNMENTS

RESULT 1  
US-09-401-476-2  
; Sequence 2, Application US/09401476  
; Patent No. 6168936  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Huaming  
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes  
; FILE REFERENCE: GC584  
; CURRENT APPLICATION NUMBER: US/09/401,476  
; CURRENT FILING DATE: 1999-09-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 583  
; TYPE: PRT  
; ORGANISM: Stachybotrys chartarum  
US-09-401-476-2

Query Match	100.0%	Score 3114;	DB 4;	Length 583;
Best Local Similarity	100.0%	Pred. No. 3.3e+290;		
Matches 583;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MISQATGAVALGLAVIGSSVDARSVAGRSTDMPSGLTRKOTOLSPPLALYEYPLPIPL 60		
DB	1	MISQATGAVALGLAVIGSSVDARSVAGRSTDMPSGLTRKOTOLSPPLALYEYPLPIPL 60		
QY	61	KAPNTVPNNTGEDIITYEMETIRPESHQIYPLLEPANMVGYSMPGPTIIVRGTESVY 120		
DB	61	KAPNTVPNNTGEDIITYEMETIRPESHQIYPLLEPANMVGYSMPGPTIIVRGTESVY 120		
QY	121	REFNSEENTSPNSVHLHGFSRAPFDGMAEDTTOPEYKDYYPNROAARMLYHDHAMS 180		
DB	121	REFNSEENTSPNSVHLHGFSRAPFDGMAEDTTOPEYKDYYPNROAARMLYHDHAMS 180		
QY	181	ITAENAYMOAGYMIQDPAEDALNLPSCYGEFDIPLVLTAKRYNADGTLFTNGEYVSF 240		
DB	181	ITAENAYMOAGYMIQDPAEDALNLPSCYGEFDIPLVLTAKRYNADGTLFTNGEYVSF 240		
QY	241	WGDVIOVNOQPMPLNVQKRTKFRFLNAVSRFLYIATISDSSTRLPFOYIADGGL 300		
DB	241	WGDVIOVNOQPMPLNVQKRTKFRFLNAVSRFLYIATISDSSTRLPFOYIADGGL 300		
QY	301	LEGPVDTFLYISMAERWEVVIDESFAGOSIDIRLPGADGIGVPEPDNDKVARFVY 360		
DB	301	LEGPVDTFLYISMAERWEVVIDESFAGOSIDIRLPGADGIGVPEPDNDKVARFVY 360		
QY	361	DEVLESPTSEVPANLRDVPFPEGGMNDPANPTDDEFTFGGRANGWTNGVTFSDYENR 420		
DB	361	DEVLESPTSEVPANLRDVPFPEGGMNDPANPTDDEFTFGGRANGWTNGVTFSDYENR 420		

QY	421	LLRVRPRTVEIWMLENNNSGWTBPRVHIHLYDFEVLRSRARGCEPEAAGLKDYVLMR	480
Db	421	LLRVRPRTVEIWMLENNNSGWTBPRVHIHLYDFEVLRSRARGCEPEAAGLKDYVLMR	480
QY	481	REVVYVEVHAYRPGCYVMHCHNLIHEHDHDMAAFNATVGDYGVNTETIDPMEPLMR	540
Db	481	REVVYVEVHAYRPGCYVMHCHNLIHEHDHDMAAFNATVGDYGVNTETIDPMEPLMR	540
QY	541	RPFLILGEFENGSGFSELATIDRIQEMASFNPAQAADDDAAEE	583
Db	541	RPFLILGEFENGSGFSELATIDRIQEMASFNPAQAADDDAAEE	583

```

RESULT      2
US-09-401-476-4
: Sequence 4, Application US/09401476
: Patent No. 6168936
: GENERAL INFORMATION:
: APPLICANT: Wang, Huaming
: TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
: FILE REFERENCE: GC584
: CURRENT APPLICATION NUMBER: US/09/401,476
: CURRENT FILING DATE: 1999-09-22
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 572
: TYPE: PRT
: ORGANISM: Stachybotrys Charatum
: US-09-401-476-4

```

Query Match	64.5%;	Score 2010;	DB 4;	length 572;
Best Local Similarity	65.4%;	Pred. No. 2.9e-184;		
Matches 383;	Conservative 71;	Mismatches 110;	Indels 22;	Gaps 10

QY	1	MISALGAVNVLGLAVICGSSVDARSVAGRSTDMSC--LTKROTQLSPPLALVYVPLPIPP	57
QY	1	MEKHTLGAALSL--LENSNNAVQASVPV--ETSPATGLFNRVAQISDPQYPMFVPIPIPP	57
QY	60	LKAPN-IVPMPNTGEDIILYYEMELRPSPHQIYDPLEBANVAVGDGMSPGPTIIVPGRGTS	118
QY	58	VKQRLVTVNPNVNGQELWYVELEKKPTHTQYVYDLDGADLVGDGMSPGPTIVPGVET	117
QY	119	VYRFVNSGENTSPNSVHLHGSFRAPFDGMAEDTTPGEXKDYVYDNRQAARMLWYDHA	178
Db	118	VYREINNAE--AENSVALHGSFRAAFDDGMAEDITEPGSKDYVYDNRQAARMLWYDHA	175
QY	179	MSITAEANVAGQGVYMIQDPAEDALNLPSCGYGEFIDPLVLTAKRNADGTLESTNGEVS	238
Db	176	MHTAEANVAGQGLVYMLTDPAEDALNLPSCGYGEFIDPLVLTAKRNADGTLESTNGEIN	235
QY	239	SFMGDVIVQVNGQPMPLNTVQPRKYPREFLAAASRSPALVLAESSENRLLPPOVYLAAG	298
Db	236	SFMGDVIVHVGQPMPEFNVPKRYRREFLDANVSRSEGLYFADTDLIDRLRPKYVASUS	295
QY	299	GLLEGVYDITLYISMAERWEVYIDESTFAGOSIDIRNLPGA--DGLVPEPFDNDKVMR	357
Db	296	GLLEHPADTSLYISMAEREYEVFEDPSDAGKTELNRNGSGTGGITDIDVNDNDKVMR	355
QY	358	FVYDDEVLESDDTSEVPANLNDVPPREGGMNDPANIIDD--ETFTSEGRANGQWTINGCVTS	415
Db	356	FVYAADTTQDPTISVPANLNDVPPF-----SPTTNPPIRQFREPDPPTWTINGCVA	407
QY	416	DVENRLLRNPVPRQTVELWRLSNSNGTHTVHILHVDYFRLNST---ARGVEPYTAAGL	472
Db	408	DYQRLRLANPVGTVGERWELINAGNWTHTIHLHVDKVISSTGNNARITWYPE--SLG	466
QY	473	KDYVMLLAREVVYVEAHYAFPCGYVMILHGNLHEDHDMAAANVYLVLDGQNTAEFD	532
Db	467	KDYVMLGRETYYVEAHYAFPCGYVYHCHNLHEDHDMAAANATVLDYQGNATVAFD	526

```
QY      533 PMEDLMPRPPELLGEFFENGSGGFSELAITDRIOEMASFNPYAQQDD 5788
      ||| ||: ||: ||||: || || ||: |||| ||: ||| ||:
Db      527 PMEELMQARPYELGEFQAQSGSFVQAVTERIQTMAYEYRPPYAADE 5722
```

RESULT 3  
 US-08-749-882A-2  
 Sequence 2, Application US/08749882A  
 Patent No. 5750388  
 GENERAL INFORMATION:  
 APPLICANT: Berka, Randy  
 APPLICANT: Thompson, Sheryl  
 APPLICANT: Xu, Feng  
 TITLE OF INVENTION: Purified Scytalidium Laccases  
 TITLE OF INVENTION: And Nucleic Acids Encoding Same  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 57503880 No. 5750388disk Of No. 5750388th America, Inc.  
 STREET: 405 Lexington Avenue - 64th Fl.  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10174  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/749,882A  
 FILING DATE: 15-NOV-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lambiris, Elias J  
 REGISTRATION NUMBER: 33,728  
 REFERENCE/DOCKET NUMBER: 4166,020-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-878-9652  
 TELEFAX: 212-878-9655  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 616 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE: Scytalidium thermophilum  
 US-08-749-882A-2

Query Match	8.28;	Score 254.5;	DB 1;	Length 616;
Best Local Similarity	25.28;	Pred. No. 9.8e-16;		
Matches 137;	Conservative 71;	Mismatches 171;	Indels 165;	Gaps 33;

[illegible]

```

Db 288 LA-----KHTMVIAD-----WVPMAMVDSLFMAVGQRYDVTIDASQAVGNWFNI 336
Qy 326 TFGAGS-----IDIRNLBGA-DGLSTGPEPEFDMTDKMRVNDVEULESPDTSEVPA 374
Db 337 TFGGQKCGKGFESHNPAPAAIFRYEQAAPDLPPTDPAAPKD-----HQCIDTIDL--PV 387
Qy 375 NLRDVP-----PPEGNMDPANPTDDE-----TTFEGRANGQWINGTFSDE----- 418
Db 388 VQKNVPYVGFEKKEGNTLPVTLHVDQAAAPHVFT-----WKING--SAADVDMDBREYLE 439
Qy 419 ---NRLLBNVP--RDYVEI-----WRLENNNSNG--WTHPVHILNDFRILSR-- 455
Db 440 YVMANDLSSITFYKNNIYKAVGDVENVMTYMLVLENDGRLSLRHPHMLHGHFETVYGRSDV 499
Qy 460 -----TARGVEYEAGLKDVMYLMARREVVYVEEAHYAEPFGVYMLHCH 502
Db 500 SPDSETRFVDPAPVDLPRLRGHNFPVR-----RDVTMLPARGWLTL-APRTDNGAMLEFCH 554
Qy 503 NLIH 506
Db 555 IAWH 558

```

RESULT 4  
 US-08-539-134-2  
 : Sequence 2, Application US/08539134  
 : Patent No. 5843745  
 : GENERAL INFORMATION:  
 : APPLICANT: Berka, Randy  
 : APPLICANT: Thompson, Sheryl  
 : APPLICANT: Xu, Feng  
 : TITLE OF INVENTION: Purified Scytalidium laccases  
 : TITLE OF INVENTION: And Nucleic Acids Encoding Same  
 : NUMBER OF SEQUENCES: 9  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: No. 58437450 No. 5843745disk Of No. 5843745th America, Inc.  
 : STREET: 405 Lexington Avenue - 64th Fl.  
 : CITY: New York  
 : STATE: NY  
 : COUNTRY: USA  
 : ZIP: 10174  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette  
 : COMPUTER: IBM Compatible  
 : OPERATING SYSTEM: DOS  
 : SOFTWARE: FastSeq for Windows Version 2.0  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/539,134  
 : FILING DATE: 04-OCT-1995  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Lambdiris, Elias J  
 : REGISTRATION NUMBER: 33,728  
 : REFERENCE/DOCKET NUMBER: 4186.010-US  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 212-878-9652  
 : TELEFAX: 212-878-9655  
 : INFORMATION FOR SEQ ID NO: 2:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 616 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : ORIGINAL SOURCE: Scytalidium thermophilum  
 : ORGANISM: Scytalidium thermophilum  
 : US-08-539-134-2

Query Match	8.28;	Score 254.5;	DB 2;	Length 616;	..
Best Local Similarity	25.28;	Pred. No. 9.8e-16;			
Matches 137;	Conservative 71;	Mismatches 171;	Indels 165;	Gaps 33;	

[illegible]

RESULT 5  
US-08-991-531-2  
Sequence 2, Application US/08991531  
Patent No. 5925554  
GENERAL INFORMATION:  
APPLICANT: Pedersen, Anders Hjelholt  
APPLICANT: Svendsen, Allan  
APPLICANT: Schneider, Palle  
APPLICANT: Rasmussen, Grethe  
APPLICANT: Cherry, Joel  
TITLE OF INVENTION: Myceliophthora And Scytalidium Laccase  
TITLE OF INVENTION: Variants  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: No. 5925554o No. 5925554dsk of No. 5925554th America, Inc  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,531  
FILING DATE: 16-DEC-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rozek, Carol  
REGISTRATION NUMBER: 36,993

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rozek, Carol  
REGISTRATION NUMBER: 36,993

REFERENCE/DOCKET NUMBER: 5125.200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 616 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-991-531-2

Query Match 8.1%; Score 253.5; DB 2; Length 616;  
Best Local Similarity 25.2%; Pred. No. 1.2e-15;  
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDILYEMERPFSSHOIYDLEPANNVGYDGMSP-----GPTIIVRGTE 117  
DB 74 PNTG-----VYRRYTPDI---TEVDNRPDGVYKEKMLINDKLGPYFANWG-D 121  
QY 118 SVYRFVNGENTSPNSVHLHGSFRRAP-----FDGNAEDTQPEYKDYYPNRQARML 172  
DB 122 TIEVYVNNHRLRTGTSIHMHGLHOKGTNYHDGANGVTECPYPPGSGRVSFRARQYGS- 180  
QY 173 WYDHAMSTIAENAYGAGVYMIODPAEDALNPSGYGFEDIP-LVTRAKRY---NADG 228  
DB 181 WYHSH-FSAQYGN---GVSGAIOINGPA---SLP-----YDIDGLVPLXDMYYSADQ 227  
QY 229 TLESTNGEVSSEFGDYIOVNGQ-----PWPMLNVP-RKYRFRFLAAVSRSFALY 278  
DB 228 LVETETLXKGAPPSDVNLINGTAKHPTTGEYAIYKLTDPDKRHLRLINMSVENHFOVS 287  
QY 279 LATSSESETRLPFOVIAADGLEGVD---TDTLYISMAERWEVVIDFS----- 325  
DB 288 LA-----KHTMTVIAAD---MVPVNAVTVSLFMAVGORYDVTIDASQAVGNWFNT 374  
QY 326 TFGAGS-----IDIRNLPGA-DGLGVEPEPNTDKVMRFVVDVLESPDTSEVPA 374  
DB 337 TFGGQKCGFSHPAPALFREGARDALPTDGAAPKD-----HQCIDTLDLS--PV 387  
QY 375 NLRDVP---FPEGGMWDANPTDDE---TTFGRANGQMTINGVTSDEVE----- 418  
DB 388 VOKNVPVDGFKRPGNTLPVTLHVDAQAARHYVT-----WKING-SAADVWMDRPLYE 439  
QY 419 ---NRLLRNP--RDYVEI-----WRLSNSNG---WTHPVHILVDFRVLSS--- 459  
DB 440 YVANNLSSIPVKNNTIVRDGVNEMTYWLVENDPEGRLSLPHMHHLGHDFEVLGRSPDV 499  
QY 460 -----TARGVEPYEAGLKDYVWLARREVYVEAHYAPFPGVYMLHCH 502  
DB 500 SPDSETRFVFPDPAVDLPRLRGNHVPV---RDVTMLPARQWMLL-AFRDNPQAWLFFCH 554  
QY 503 NLH 506  
DB 555 IAXH 558

RESULT 6  
US-09-032-315-9  
Sequence 9, Application US/09032315  
Patent No. 5985818  
GENERAL INFORMATION:  
APPLICANT: Svendsen, Allan  
TITLE OF INVENTION: LACCASE MUTANTS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 59858180 No. 5985818disk of No. 5985818ch America, Inc.  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY

COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/032,315  
FILING DATE: 27-FEB-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Rozek, Carol  
REGISTRATION NUMBER: 36,993  
REFERENCE/DOCKET NUMBER: 5200.200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 616 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-032-315-9

Query Match 8.1%; Score 253.5; DB 2; Length 616;  
Best Local Similarity 25.2%; Pred. No. 1.2e-15;  
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDILYEMERPFSSHOIYDLEPANNVGYDGMSP-----GPTIIVRGTE 117  
DB 74 PNTG-----VYRRYTPDI---TEVDNRPDGVYKEKMLINDKLGPYFANWG-D 121  
QY 118 SVYRFVNGENTSPNSVHLHGSFRRAP-----FDGNAEDTQPEYKDYYPNRQARML 172  
DB 122 TIEVYVNNHRLRTGTSIHMHGLHOKGTNYHDGANGVTECPYPPGSGRVSFRARQYGS- 180  
QY 173 WYDHAMSTIAENAYGAGVYMIODPAEDALNPSGYGFEDIP-LVTRAKRY---NADG 228  
DB 181 WYHSH-FSAQYGN---GVSGAIOINGPA---SLP-----YDIDGLVPLXDMYYSADQ 227  
QY 229 TLESTNGEVSSEFGDYIOVNGQ-----PWPMLNVP-RKYRFRFLAAVSRSFALY 278  
DB 228 LVETETLXKGAPPSDVNLINGTAKHPTTGEYAIYKLTDPDKRHLRLINMSVENHFOVS 287  
QY 279 LATSSESETRLPFOVIAADGLEGVD---TDTLYISMAERWEVVIDFS----- 325  
DB 288 LA-----KHTMTVIAAD---MVPVNAVTVSLFMAVGORYDVTIDASQAVGNWFNT 374  
QY 326 TFGAGS-----IDIRNLPGA-DGLGVEPEPNTDKVMRFVVDVLESPDTSEVPA 374  
DB 337 TFGGQKCGFSHPAPALFREGARDALPTDGAAPKD-----HQCIDTLDLS--PV 387  
QY 375 NLRDVP---FPEGGMWDANPTDDE---TTFGRANGQMTINGVTSDEVE----- 418  
DB 388 VOKNVPVDGFKRPGNTLPVTLHVDAQAARHYVT-----WKING-SAADVWMDRPLYE 439  
QY 419 ---NRLLRNP--RDYVEI-----WRLSNSNG---WTHPVHILVDFRVLSS--- 459  
DB 440 YVANNLSSIPVKNNTIVRDGVNEMTYWLVENDPEGRLSLPHMHHLGHDFEVLGRSPDV 499  
QY 460 -----TARGVEPYEAGLKDYVWLARREVYVEAHYAPFPGVYMLHCH 502  
DB 500 SPDSETRFVFPDPAVDLPRLRGNHVPV---RDVTMLPARQWMLL-AFRDNPQAWLFFCH 554  
QY 503 NLH 506  
DB 555 IAXH 558





APPLICATION NUMBER: US/09/032,315  
FILING DATE: 27-FEB-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Rozek, Carol  
REGISTRATION NUMBER: 36,993  
REFERENCE/DOCKET NUMBER: 5200.200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ. ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 616 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-396-260-9

Query Match 8.1%; Score 253.5; DB 4; Length 616;  
Best Local Similarity 25.2%; Pred. No. 1,2e-15;  
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDIYEMETIRPESHOIYDLEPANNVGYDGMSPP-----GPTIIVRGTE 117  
DB 74 PNTG-----VYRRYTFDI---TEVDNRPGRPDGVYKERLMLINDKLGLPTVFANMG-D 121  
QY 118 SVYRFVNGSGENTSPNSVYHLHGSFSRAP-----FDGNAEDTTPQGEYKDYVPNNRQAARML 172  
DB 122 TIEYTVNNHLRTGTSTHMHGKHOKGTNYHDGANGVTECPYIPGGSRYVSFRARQYCTS- 180  
QY 173 WYDHAMSTIAENAYMOAGVYMIODPAEDALNLPSSYGFEDIPL-VLTAKRY---NADG 228  
DB 181 WYHSH-FSAQYGN---GVSGAIOINGPA---SLP-----YDIDGLVPLXDMYKXSADQ 227  
QY 229 TLFSTNGEVSSFGVDYIQVNGQ-----PWPMLNVQP-RKYRREFELNAVSRSFALY 278  
DB 228 LVETETLKGNAPPSDNLVINGTAKHPTTGEGEYAIYVLTPTDKRRLRLIMSVENHFOVS 287  
QY 279 LATESESETRLPFOVYIADGGLLEGVD---TPTLYISMAERWEVVIDFS----- 325  
DB 288 LA-----KHTMTVLIAD---MVPVNAVTVDLSFMAVQORYDVTIDASQAVGNWYFNI 336  
QY 326 TFGAQS-----IDIRNLPGA-DGLGVEPEFNDTKVMRFVYVDEVLESPTSEVPA 374  
DB 337 TFGGQKCGFSHPNAPARAIIRYEGAPDALPTDGAAPKD-----HQCIDTIDLSS--PV 387  
QY 375 NLADVP---FPEGKMDPANTPTDDE---TTFGRANGQWITNGVTSQVDE----- 418  
DB 388 VQKNVPYDGFVKBERGNTLPYTLHVDAQAAAPHVFT-----WKING-SAADVDMDRVLE 439  
QY 419 ---NRLRNVP---RDYVEI-----WRLENNNSNG---WTHPVHILVDFRVLRSR--- 459  
DB 440 YVMNNDLSSIPVKNNTIVRYDGVNEMTYWLVENDEBGLSLPHEPHMLGHDFVYLGSRSDV 499  
QY 460 -----TARGVERTEAAGLKDYYMLARREYVYVEAHYAPFGVYMLHCH 502  
DB 500 SPDSETRFVDPADVLPRLRGHNPVR---RDYTMLPARGWMLL-ARTDNPQAWLHFCH 554  
QY 503 NLIH 506  
DB 555 IAXH 558

RESULT 11  
US-09-518-901-2  
Sequence 2, Application US/09518901  
Patent No. 6218170  
GENERAL INFORMATION:  
APPLICANT: Svendsen, Allan  
TITLE OF INVENTION: LACCASE MUTANTS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:

ADDRESSEE: NO. 6218170d NO. 6218170disk of NO. 6218170th America  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/518,901  
FILING DATE: 06-Mar-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/028,887  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Gregg, Valeta A.  
REGISTRATION NUMBER: 35,127  
REFERENCE/DOCKET NUMBER: 5201.200-US  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 616 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-518-901-2

Query Match 8.1%; Score 253.5; DB 4; Length 616;  
Best Local Similarity 25.2%; Pred. No. 1,2e-15;  
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDIYEMETIRPESHOIYDLEPANNVGYDGMSPP-----GPTIIVRGTE 117  
DB 74 PNTG-----VYRRYTFDI---TEVDNRPGRPDGVYKERLMLINDKLGLPTVFANMG-D 121  
QY 118 SVYRFVNGSGENTSPNSVYHLHGSFSRAP-----FDGNAEDTTPQGEYKDYVPNNRQAARML 172  
DB 122 TIEYTVNNHLRTGTSTHMHGKHOKGTNYHDGANGVTECPYIPGGSRYVSFRARQYCTS- 180  
QY 173 WYDHAMSTIAENAYMOAGVYMIODPAEDALNLPSSYGFEDIPL-VLTAKRY---NADG 228  
DB 181 WYHSH-FSAQYGN---GVSGAIOINGPA---SLP-----YDIDGLVPLXDMYKXSADQ 227  
QY 229 TLFSTNGEVSSFGVDYIQVNGQ-----PWPMLNVQP-RKYRREFELNAVSRSFALY 278  
DB 228 LVETETLKGNAPPSDNLVINGTAKHPTTGEGEYAIYVLTPTDKRRLRLIMSVENHFOVS 287  
QY 279 LATESESETRLPFOVYIADGGLLEGVD---TPTLYISMAERWEVVIDFS----- 325  
DB 288 LA-----KHTMTVLIAD---MVPVNAVTVDLSFMAVQORYDVTIDASQAVGNWYFNI 336  
QY 326 TFGAQS-----IDIRNLPGA-DGLGVEPEFNDTKVMRFVYVDEVLESPTSEVPA 374  
DB 337 TFGGQKCGFSHPNAPARAIIRYEGAPDALPTDGAAPKD-----HQCIDTIDLSS--PV 387  
QY 375 NLADVP---FPEGKMDPANTPTDDE---TTFGRANGQWITNGVTSQVDE----- 418  
DB 388 VQKNVPYDGFVKBERGNTLPYTLHVDAQAAAPHVFT-----WKING-SAADVDMDRVLE 439  
QY 419 ---NRLRNVP---RDYVEI-----WRLENNNSNG---WTHPVHILVDFRVLRSR--- 459  
DB 440 YVMNNDLSSIPVKNNTIVRYDGVNEMTYWLVENDEBGLSLPHEPHMLGHDFVYLGSRSDV 499  
QY 460 -----TARGVERTEAAGLKDYYMLARREYVYVEAHYAPFGVYMLHCH 502  
DB 500 SPDSETRFVDPADVLPRLRGHNPVR---RDYTMLPARGWMLL-ARTDNPQAWLHFCH 554

QY 503 NLH 506  
DB 555 IAXH 558

RESULT 12  
US-09-576-281-9  
; Sequence 9, Application US/09576281  
; Patent No. 6277611  
; GENERAL INFORMATION:  
; APPLICANT: Pedersen, Anders  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Schneider, Palle  
; APPLICANT: Rasmussen, Grethe  
; APPLICANT: Cherry, Joel  
; TITLE OF INVENTION: LACCASE MUTANTS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 62776110 No. 6277611disk of No. 6277611th America  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/576,281  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/993,318  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Greig, Valeta A.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 5032.200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 616 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-576-281-9

Query Match 8.1%; Score 253.5; DB 4; Length 616;  
Best Local Similarity 25.2%; Pted. No. 1.2e-15;  
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDIYEMIEIPFSHOIYDLEBANVGYDGMSPP-----GPTIIVPGTE 117  
DB 74 PNTG-----VYRRTPTI---TEVDNRPFGDGYIKELMLINKLLGPTFANMG-D 121  
QY 118 SVYRVNNGENTSPNVHLHGSFSRAP-----FDGMAEDTTPGEGKDYVYVNRQAARML 142  
DB 122 TIEVTANNHRTNGTSTHMHGLHKGFTNYHDGANGVTECP1PPGGSFVYSFRARQVGTGTS- 180  
QY 173 WYDHDAMSTIAENAYMGQAGVYMIODPAEDALMLPSGYGEFDPL-VLTAKRY---NADG 228  
DB 181 WYSH-FSAQYGN---GVSQAIQINGPA---SLP-----YDIDGLVLPXDMVYKSAQD 227  
QY 229 TLFSTNGEVSFPQGVYQVNGQ-----PWPMLNVQP-RKYRFRFLNAAVSRSPALY 278  
DB 228 LVLETTLAKGNAPSDNVLLINGTAKHPTTGGEYAIYKLTDPKRRRLRLIMMSVENHFOVS 287

QY 279 LATSEDESTRLPFOVIAADGGLLEGPD---TDTLYISAERNEVIDFS----- 325  
DB 288 LA-----KHTMTVIAAD-----MVPNAMTVDSLEPMAYGQRYDVTIDASQAVGNWYNI 336  
QY 326 TFGAQS-----IDIRNLPGA--DGLGVEPEPNTDKVMRFVYDEVLESDTSEVA 374  
DB 337 TFGGQCKGCFSHNPAPAAIFREYGADALPTDGAAPKD-----HGLDIDLIS--PV 387  
QY 375 NLRDVP-----FPEGWMDPANPTDDE-----PFTFGRANGQWITNGVTSQVY 418  
DB 388 VQKNVPYDGFVKPEGNTLPYTLHAVDQAAPHVTT-----WKING-SAADDVMDRPLYE 439  
QY 419 ---NRLRNVP--RDVVEI-----WRLSNSNG---WTHPVHILVDFYLSRS--- 459  
DB 440 YVANNLSSIPVKNINIVRDGVNEMWTYLVENDPEGRLSLPHRMHLHGHDFFVLGRSPDV 499  
QY 460 -----TARGVEPYEAGLKQVYLARREVVYEAHYAPFQVYVYLLHCH 502  
DB 500 SPDSETRFVFPADVLPRLGNHVPV---RDVTMLPARQWLL--AFRDNPGAWLFFCH 554  
QY 503 NLH 506  
DB 555 IAXH 558

RESULT 13  
US-08-940-661A-2  
; Sequence 2, Application US/08940661A  
; Patent No. 5795760  
; GENERAL INFORMATION:  
; APPLICANT: BERKA, Randy Michael  
; APPLICANT: BROWN, Stephen H.  
; APPLICANT: XU, Feng  
; APPLICANT: SCHNEIDER, Palle  
; APPLICANT: OXENB LL, Karen M.  
; APPLICANT: ASLING, Dorrit A.  
; TITLE OF INVENTION: PURIFIED MYCELIOPHTHORA LACCASES AND NUCLEIC  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 57957600 No. 5795760disk of No. 5795760th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/940,661A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/441,146  
; FILING DATE: 15-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harrington, James J.  
; REGISTRATION NUMBER: 38,711  
; REFERENCE/DOCKET NUMBER: 4184.010-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 620 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-940-661A-2







000000

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:16 ; Search time 58.78 Seconds  
(without alignments)  
1101.668 Million cell updates/sec

Title: US-09-656-640A-2

Perfect score: 3114

Sequence: 1 MISQALGVALGLAVIGSS.....IQEMASFPYQAQDDAAEE 583

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
A.Geneseq\_032802:\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3114	100.0	583	22	Stachybotrys chart
2	3114	100.0	583	22	Stachybotrys chart
3	3090	99.2	578	22	Bilirubin oxidase.
4	2010	64.5	572	14	Bilirubin oxidase.
5	1829	58.7	602	21	Amino acid sequenc
6	1762	56.6	627	21	Bipolaris spicifera
7	1762	56.6	627	21	Bipolaris spicifera
8	1723.5	55.3	627	21	Curvularia pallesc
9	1723.5	55.3	627	21	Curvularia pallesc
10	1716.5	55.1	594	20	Stachybotrys chart
11	1716.5	55.1	594	20	Stachybotrys pheno

12	1716.5	55.1	594	21	Stachybotrys chart
13	1716.5	55.1	594	21	Stachybotrys chart
14	605	19.4	474	22	Micromonospora eye
15	504	16.2	511	22	C glutamicum prote
16	498.5	16.0	497	22	Corynebacterium gl
17	454	14.6	570	21	Arabidopsis thalia
18	454	14.6	574	21	Arabidopsis thalia
19	454	14.6	581	21	Arabidopsis thalia
20	453	14.5	477	22	Arabidopsis thalia
21	355.5	11.4	410	22	S. epidermidis ope
22	342	11.0	113	21	S. epidermidis ope
23	342	11.0	113	21	Amerosporium atrum
24	313	10.1	348	22	S. epidermidis ope
25	254.5	8.2	616	18	Scytalidium thermo
26	253.5	8.1	616	19	Scytalidium thermo
27	253.5	8.1	616	19	Scytalidium thermo
28	253.5	8.1	616	19	Scytalidium thermo
29	253.5	8.1	616	19	Scytalidium thermo
30	253.5	8.1	616	21	Scytalidium thermo
31	253	8.1	620	18	Myceliophthora the
32	243	7.8	573	19	Myceliophthora the
33	241	7.7	573	19	Myceliophthora the
34	241	7.7	573	19	Myceliophthora the
35	241	7.7	573	19	Myceliophthora the
36	241	7.7	573	19	Myceliophthora the
37	241	7.7	573	19	Myceliophthora the
38	241	7.7	573	21	Myceliophthora the
39	241	7.7	620	17	Myceliophthora the
40	241	7.7	620	18	Myceliophthora the
41	241	7.7	620	21	Myceliophthora the
42	226.5	7.3	573	19	M. thermophila lac
43	226.5	7.3	573	19	Myceliophthora the
44	225	7.2	573	19	Myceliophthora the
45	225	7.2	573	19	Myceliophthora the

#### ALIGNMENTS

RESULT 1	
AA1981505	standard; Protein; 583 AA.
AA1981505;	
18-JUN-2001	(first entry)
Stachybotrys chartarum phenol oxidase B enzyme.	
Stachybotrys phenol oxidase B; spob; phenol oxidizing enzyme;	
detergent; paper production; pulp production; textile; food industry;	
bleaching.	
Stachybotrys chartarum.	
WO200121748-A1.	
29-MAR-2001.	
06-SEP-2000; 2000WO-EP08840.	
22-SEP-1999; 99EP-0203120.	
(UNITL ) UNILEVER NV.	
(UNITL ) UNILEVER PLC.	
(HIND-) HINDUSTAN LEVER LTD.	
Convents D, Doornik M, De Vries CH, Wang H;	
WPI, 2001-273462/28.	
N-PSDB; AAF82586.	
New detergent compositions comprising a phenol oxidizing enzyme useful	

PT In detergent or cleaning compositions, fiber treatment, processing,  
PT finishing or production, paper and pulp production, or in starch  
PT processing applications

XX Claim 1; Fig 2; 46pp; English.

CC The present sequence is a Stachybotrys chartarum phenol oxidizing  
CC enzyme. The invention relates to detergent compositions comprising one or  
CC more surfactants and a phenol oxidizing enzyme having at least 68%  
CC identity to the Stachybotrys chartarum phenol oxidizing enzyme. Phenol  
CC oxidizing enzymes may be used in the detergent, paper, pulp, textile and  
CC food industries. They are used for preventing the transfer of dyes in  
CC solution from one textile to another during detergent washing, or in  
CC modifying the colour associated with dyes and coloured compounds having  
CC different chemical structures, such as in pulp and paper bleaching,  
CC bleaching the colour of stains on fabric and in detergent and textile  
CC applications.

XX Sequence 583 AA;

Query Match 100.0%; Score 3114; DB 22; Length 583;  
Best Local Similarity 100.0%; Pred. No. 8.9e-275;

Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISQAIGAVAGLAVIGSSVDARSVAGRSTDMPSGLTKRQTLSPPLALYEVLPIPL 60  
DB 1 misqaigavaglavigssvdarsvagrstdmgsqtkrqtlspplalyevlpipl 60

QY 61 KAPNTVPNPTGEDIILYEKEIRPFESHQIYRPLEPANMVGSDMSGPTIIVRGRESVY 120  
DB 61 kapntvpnptgediilylemeirpfishqlypdlpanmvygdmspgtliivprgesv 120

QY 121 RFVNSGENTS PNSVHLHGSESRAPFDGWAEDTQPGEXKYDYYPNROAARMLYHHDHAMS 180  
DB 121 rfvnsentspnsvhlhgsefsrapfdgwaedtqpgeykydyypnrgaarmllyhdhams 180

QY 121 rfvnsentspnsvhlhgsefsrapfdgwaedtqpgeykydyypnrgaarmllyhdhams 180  
DB 121 rfvnsentspnsvhlhgsefsrapfdgwaedtqpgeykydyypnrgaarmllyhdhams 180

QY 181 ITAENAYMGOAGVYMIQDPAEDALNPSGYGEFDPVLVITAKRYNADGTLFSTNGEVSSF 240  
DB 181 itaenaymgaagvymiqdpedalnipsgygefdpvlvltakrynadgltfstngvssf 240

QY 241 WGDVIOVNGOPMPLNWPORRKYRFRFLNAVSRFALYLATSDSETRLPFOVIADGGL 300  
DB 241 wgdviovngopmplnvprkyrfrflnaavsrfsalylatsestetrlpfqviadgg 300

QY 301 LEGPVDTDLTLYISMAERWEVVIDFTFAGOSIDIRNLPGADGLGVPEPFDNTDKVRFVY 360  
DB 301 legpvdtdltlyismaerwevidfstfagsidirlnpgadglgvpefdntdkvmrfv 360

QY 361 DEYLESPTSEVPANLRDVPFPEGGMWDPANPTDDETFTEGRANGWTINGVTFSDVENR 420  
DB 361 devylesptsevpantldvfpfeggmwdpanptddeftegrangwtlingvtfsvdenr 420

QY 421 LLRNVRODVEITRLENNNGWTHPVHILVDBRVLSRSTARGVEPEYEAAGLTDVWMLAR 480  
DB 421 llrnvrodveitrlennngwthpvhlvdbrvlsrstargvepeyeadgldvwwmlar 480

QY 481 RIEVVYEAHAPFPYGYMLHCHNLHEDHDMMAAFNVTVLGDGYVYTERIDMEPLMRP 540  
DB 481 revvyeahapfpgyymhcnhlhedhdmmaafnvtlvgdgyvyteridmepmrp 540

QY 541 REPPLGEEFNGSGDSSELAITTDRIQMASEFNYPQAODDDAAEF 583  
DB 541 repplgeefngsgdsselaittdriqmasenypqaodddaaef 583

RESULT 2  
AAB20097  
ID AAB20097 standard; Protein; 583 AA.  
AC AAB20097;  
XX  
DT 23-APR-2001 (first entry)

XX Stachybotrys chartarum phenol oxidase B.  
DE phenol oxidizing enzyme; phenol oxidase B; spob gene; bleach;  
XX pulp; paper; textile; detergent.

XX Stachybotrys chartarum.

OS US6168936-B1.

PN 02-JAN-2001.

XX 22-SEP-1999; 9905-0401476.

XX 22-SEP-1999; 9905-0401476.

PA (GENM ) GENENCOR INT INC.

XX Wang H;

DR MPI; 2001-136715/14.

XX N-PSDB; AAF30028, AAF20029.

PT New phenol oxidizing enzyme, also useful in the detergent, paper and  
PT pulp, textile or food industries, especially in modifying the colour  
PT associated with dyes and coloured compounds, as well as in anti-dye  
PT transfer applications

PS Claim 1; Fig 2; 23pp; English.

CC The present sequence is that of Stachybotrys chartarum MUC1 38898  
CC phenol oxidase B, as deduced from isolated genomic DNA (see  
CC AAF30028). The invention provides phenol oxidizing enzymes such  
CC as phenol oxidase B, or enzymes having at least 68% identity to  
CC it, nucleic acids encoding them, expression vectors, filamentous  
CC fungus and yeast cells, and methods for the recombinant  
CC production of the phenol oxidizing enzymes. The enzymes are useful  
CC for bleaching pulp and paper, fabric stains, and in detergent and  
CC textile applications. They show optimal activity at pH range 5-11,  
CC 7-10.5 or 8-10, and at 20-60 or 20-40 degree C.

XX Sequence 583 AA;

Query Match 100.0%; Score 3114; DB 22; Length 583;  
Best Local Similarity 100.0%; Pred. No. 8.9e-275;

Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISQAIGAVAGLAVIGSSVDARSVAGRSTDMPSGLTKRQTLSPPLALYEVLPIPL 60  
DB 1 misqaigavaglavigssvdarsvagrstdmgsqtkrqtlspplalyevlpipl 60

QY 61 KAPNTVPNPTGEDIILYEKEIRPFESHQIYRPLEPANMVGSDMSGPTIIVRGRESVY 120  
DB 61 kapntvpnptgediilylemeirpfishqlypdlpanmvygdmspgtliivprgesv 120

QY 121 RFVNSGENTS PNSVHLHGSESRAPFDGWAEDTQPGEXKYDYYPNROAARMLYHHDHAMS 180  
DB 121 rfvnsentspnsvhlhgsefsrapfdgwaedtqpgeykydyypnrgaarmllyhdhams 180

QY 121 rfvnsentspnsvhlhgsefsrapfdgwaedtqpgeykydyypnrgaarmllyhdhams 180  
DB 121 rfvnsentspnsvhlhgsefsrapfdgwaedtqpgeykydyypnrgaarmllyhdhams 180

QY 181 ITAENAYMGOAGVYMIQDPAEDALNPSGYGEFDPVLVITAKRYNADGTLFSTNGEVSSF 240  
DB 181 itaenaymgaagvymiqdpedalnipsgygefdpvlvltakrynadgltfstngvssf 240

QY 241 WGDVIOVNGOPMPLNWPORRKYRFRFLNAVSRFALYLATSDSETRLPFOVIADGGL 300  
DB 241 wgdviovngopmplnvprkyrfrflnaavsrfsalylatsestetrlpfqviadgg 300

QY 301 LEGPVDTDLTLYISMAERWEVVIDFTFAGOSIDIRNLPGADGLGVPEPFDNTDKVRFVY 360  
DB 301 legpvdtdltlyismaerwevidfstfagsidirlnpgadglgvpefdntdkvmrfv 360

QY 361 DEYLESPTSEVPANLRDVPFPEGGMWDPANPTDDETFTEGRANGWTINGVTFSDVENR 420  
DB 361 devylesptsevpantldvfpfeggmwdpanptddeftegrangwtlingvtfsvdenr 420

```

|||||
Db 361 devlespdtevpnldvfppeggnwdpanplddefltgrangvclngvltfsvdenr 420
Qy 421 LLRNPBDTVEIMRLNNSNGWTHPVHILVDFRVLSRSTARGVEPEAGLKDVMVLAR 480
    |||||||
Db 421 llrvpdtvlelwrleannsgwtbphvhlvdfrvlsrstargvepeaglkdvwvlar 480
Qy 481 REVVYEAHAPFPGVYMLCHNLIHEDHDMMAFNNTVLGDYGYNTTEFTDPEPLMRP 540
    |||||||
Db 481 revvyeahapfpgvymhcnhlhedhmmaafnvtlvigdgyntefldpmeplwtrp 540
Qy 541 RPFLLGEFENGSGDFSELATDRIQEMASFNPPYAQADDDAAEE 583
    |||||||
Db 541 rpfllgefengsgdfselatdriqemasfnpyaqaddaaee 583

RESULT 3
AAB81506 standard; Protein: 578 AA.
ID AAB81506;
XX
AC AAB81506;
XX
DT 18-JUN-2001 (first entry)
XX
DE Billrubin oxidase.
XX
KW Billrubin oxidase; phenol oxidizing enzyme; phenol oxidase B;
KM detergent; paper production; pulp production; textile; food industry;
KW bleaching.
XX
OS unidentified.
XX
PN WO200121748-A1.
XX
PD 29-MAR-2001.
XX
PE 06-SEP-2000; 2000WO-EP08840.
XX
PR 22-SEP-1999; 99EP-0203120.
XX
PA (UNITL ) UNILEVER NV.
PA (UNITL ) UNILEVER PLC.
PA (HIND-) HINDUSTAN LEVER LTD.
XX
PI Convents D, Doornik M, De Vries CH, Wang H;
XX
DR MPI; 2001-273462/28.
XX
PT New detergent compositions comprising a phenol oxidizing enzyme useful
PT in detergent or cleaning compositions, fiber treatment, processing,
PT finishing or production, paper and pulp production, or in starch
PT processing applications
XX
PS Example 4; Fig 4; 46pp; English.
XX
CC The present sequence was used for comparison with the stachybotrys
CC charitum phenol oxidizing enzyme. The invention relates to detergent
CC compositions comprising one or more surfactants and a phenol oxidizing
CC enzyme having at least 68% identity to the stachybotrys charitum phenol
CC oxidizing enzyme. Phenol oxidizing enzymes may be used in the detergent,
CC paper, pulp, textile and food industries. They are used for preventing
CC the transfer of dyes in solution from one textile to another during
CC detergent washing, or in modifying the colour associated with dyes and
CC coloured compounds having different chemical structures, such as in pulp
CC and paper bleaching, bleaching the colour of stains on fabric and in
CC detergent and textile applications.
XX
SQ Sequence 578 AA:

```

Query Match 99.2%; Score 3090; DB 22: Length 578;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-272;  
 Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MISQIGAVALLGLAVIGSSVDARSVAGRSNDMPSGLTTRKROTSPPLALYEVPPIPL 60
    |||||||
Db 1 misqigavallglavigssvdarsvagrscdmpsgltkrtqlspplalyevprippl 60
Qy 61 KAPTVPNPTGEDILTYEMEIRPFSSHQIYPDLEPAMWGDGSPPTIIVRGTSVV 120
    |||||||
Db 61 kaptvpnpntgelltyemairpfshqiyppdleparmwgdgmsppllivrgtssvv 120
Qy 121 RFVNSGNTSPNSVHLHGSSSRAPFDGWAEDTQPGCYKYYPNROAARMLWHDHAMS 180
    |||||||
Db 121 rfvnsgntspnsvhlhgsstsrappfdgwaedtlqpgcykyypnrqaarmlwyhdhams 180
Qy 181 ITAENAYMGAGVYMIODPDALNIPSGGEPIPLVLAKRKNACGLFSTNGEYSSF 240
    |||||||
Db 181 itaenaymgagvymiodpdalnipsygetdplvlakrknadgclfstngevssf 240
Qy 241 WGDVYQVNGQPPWMLANQPKRYRFRFLNAAVSRSFALYLATSEDSERLPPFOVIAADGGL 300
    |||||||
Db 241 wgdvlyqngqppwmlnvqprkkyrrfrlnaavsrfsalylatsedserlppfyiaadggl 300
Qy 301 LEGPVDITDLYISAEKWEVYIDFSTFAGSIDIRNLPGADGLGVEPDNTDKWRFV 360
    |||||||
Db 301 legpvditdlyismaerwvvidfstfagsidirnlpagdglyvepefdntdkwrfv 360
Qy 361 DEVLESPDTSEVPANLDRVFPPEGGNWDPANPTDDEFTEFRANGQWTINGVFEVDENR 420
    |||||||
Db 361 devlespdtevpnldvfppeggnwdpanptddefltgrangvclngvltfsvdenr 420
Qy 421 LLRNPBDTVEIMRLNNSNGWTHPVHILVDFRVLSRSTARGVEPEAGLKDVMVLAR 480
    |||||||
Db 421 llrvpdtvlelwrleannsgwtbphvhlvdfrvlsrstargvepeaglkdvwvlar 480
Qy 481 REVVYEAHAPFPGVYMLCHNLIHEDHDMMAFNNTVLGDYGYNTTEFTDPEPLMRP 540
    |||||||
Db 481 revvyeahapfpgvymhcnhlhedhmmaafnvtlvigdgyntefldpmeplwtrp 540
Qy 541 RPFLLGEFENGSGDFSELATDRIQEMASFNPPYAQADDD 578
    |||||||
Db 541 rpfllgefengsgdfselatdriqemasfnpyaqadd 578

```

```

RESULT 4
AAR40843 standard; Protein: 572 AA.
ID AAR40843;
XX
AC AAR40843;
XX
DT 24-FEB-1994 (first entry)
XX
DE Billrubin oxidase.
XX
KW Billrubin; oxidase; analytical; BO: expression vector; PCR;
KW polymerase chain reaction.
XX
OS Myrothecium verrucaria.
XX
FH key
FH Reptide 1..38
FT /label= sig_peptide
FT Protein 39..534
FT /label= mat_protein
PN JP05199882-A.
XX
PD 10-AUG-1993.
XX
PF 24-JAN-1992; 92JP-0034126.
XX
PR 24-JAN-1992; 92JP-0034126.
XX
PA (AMAN ) AMANO PHARM KK.
XX

```

DR WPI; 1993-284681/36.  
DR N-PSDB; AAQ47790.  
XX Bilirubin oxidase prep. useful as an analytical enzyme - by  
XX culturing bilirubin oxidase in transformant culture  
XX Claim 1; Page 29-32; 32pp; Japanese.  
PS The sequence encodes bilirubin oxidase. The protein produced has a  
CC 38 amino acid signal peptide which is removed to give the mature  
CC protein (AA040843).  
XX  
SQ Sequence 572 AA;

Query Match 64.5%; Score 2010; DB 14; Length 572;  
Best Local Similarity 65.4%; Pred. No. 3.7e-174;  
Matches 303; Conservative 71; Mismatches 110; Indels 22; Gaps 10;

QY 1 MISQAIGAVAGLAVIGSSVDARSVAGRSTDMPSG-LTRKROTLSPLALYEVLPPIPP 59  
D 1 mfhltgaaalsl-lfnsnaevsyp--etspatghlfrvqalspqymfvcplpipp 57  
XX  
QY 60 LKAPN-TVPNPTGEDIITYEMEIRPFSHQIYFDLEPANWGDGMSGPTIIVPGSTES 118  
D 58 vkqprltvtpvnggelwlyveveikpfbhqvypdlgsadivgydgmsspgtfgvprgvev 117  
XX  
QY 119 VYFVNSGENTSPKNSVHLHGSFRRAPDDGAEEDTQGEKDYKYYPRQAAKRLWYHDI 338  
D 118 vyfllmae--apnsyllhgsfrraafdgwaedltspgskdyyprrgsartclwyhdha 135  
XX  
QY 179 MSITAENAYMGAGVYMIODPADALNPSGYGEFDIPLVLTAKRYNADGTFSTNGEVS 238  
D 176 mhtaenayrgagqlymltdpdaedlnlpsgygefdlmltskylanganlvtngeln 235  
XX  
QY 239 SFWDYIQVNGDWMKLNKOPKRYRREFLNAVSRSAVLALASDSEETLPPQVIAADG 238  
D 236 sfwgdvlyhngvgwprknpvprkyrriflidaavsrfglyadtdaltrlpikviasds 235  
XX  
QY 299 GLEGEVDVDTLYISMAERWEVVIDSTFAGOSIDIRNLPGA-DGLGVEPEFNTDKVMR 357  
D 296 gllhepadscllyismaeyevfdfsdygsktelrnlgsyggjcdydnctdkymr 355  
XX  
QY 358 FVVDVLESDPTSEVPANLRDVPFPEGGMNDPNDTD--ETFTFGRANGQMTINGVTS 415  
D 356 fvaadttcgtdevsypanlrtdvpf-----sptlntrpgrfgrtqgtclwngaf 407  
XX  
QY 416 DVENRLRNVRPRTVEIWRLENNNGMTWHPVHILVDFRLSRST--ARGVEPYEAGL 472  
D 408 dvgnrllanvpvgltveawelnaagngwclphlhdvfkvlsrtsgnatltmpe-sgl 466  
XX  
QY 473 KDVVNLAREVVYVEAHYAPFPGVYMLHCHNLTHEDHMMAAFNVLADGYNYTEFID 532  
D 467 kdvvnlgrrctevveahyepfpgvymfchcnllhedhmmaafnatvlpolygnatvfv 526  
XX  
QY 533 PNEPLMRPRPFLIGEFENGSGDFSELAIDRIEMASFNPNYQADD 578  
D 527 pmeelwgarpyelgefegagsgfsvgavterlqtmayrxyraade 572  
XX  
RESULT 5  
AA69204  
ID AA69204 standard; Protein; 602 AA.  
XX  
AC AA69204;  
XX  
DT 30-MAY-2000 (first entry)  
XX  
DE Amino acid sequence of a phenol oxidizing enzyme.  
XX  
XX Phenol oxidizing enzyme; fungus; redox reaction; detergent; bleaching;  
KW fabric; pulp; paper; decolourisation; plant-derived food product;  
KW coloured compound; porphyrin; tannin; polyphenol; carotenoid;  
KW

KW anthocyanin; Mallard reaction product.  
XX  
XX Acremonium murorum.  
OS  
XX WO200005349-A1.  
PN  
XX  
XX 03-FEB-2000.  
PD  
XX  
XX 13-JUL-1999; 99WO-EP04922.  
PF  
XX  
XX 21-JUL-1998; 98EP-0202454.  
PR  
XX  
XX (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
PA (HIND-) HINDUSTAN LEVER LTD.  
XX  
PI Convents D, Goka RJ, Van Der Heiden M, Swarthoff T, Verrips CT;  
DR  
XX  
XX WPI; 2000-195101/17.  
DR N-PSDB; AA61243.  
XX  
XX Phenol-oxidizing enzyme from Acremonium, used in detergent compositions  
PT for bleaching stains on fabrics -  
XX  
XX Claim 5; Page 41-43; 45pp; English.  
XX  
XX The present sequence represents a phenol oxidizing enzyme from the  
CC fungus Acremonium murorum. The enzyme has the CBS accession number  
CC 157.72. The enzyme catalyses redox reactions and is specific for  
CC molecular oxygen as the electron acceptor. The phenol oxidizing enzyme  
CC is specifically used in detergents for bleaching stains on fabrics,  
CC but also for bleaching pulp and paper and for decolourisation of  
CC plant-derived food products. The enzyme has a pH optimum in the  
CC alkaline to neutral range and can bleach a wide variety of coloured  
CC compounds, e.g. porphyrins, tannins, polyphenols, carotenoids,  
CC anthocyanins and Mallard reaction products.  
XX  
SQ Sequence 602 AA;

Query Match 58.7%; Score 1829; DB 21; Length 602;  
Best Local Similarity 58.7%; Pred. No. 1.3e-157;  
Matches 352; Conservative 83; Mismatches 131; Indels 34; Gaps 7;

QY 5 AIGAVAGLAVIGSSV-----DARSVAGRSTDMPSGLTRKROTLS 45  
D 6 alrelalylslkgaqampkfelidpeeaaalaavedpandlqrrspndlqrrspls 65  
XX  
QY 46 PPLALYEVLPITPLKAP-NTVPNPTGEDIITYEMEIRPFSHQIYFDLEPANWGDGM 104  
D 66 payllfqpalslppvkeplftvnpnyngeldyvelkfhfseqvfpdlqpadlvgydgl 125  
XX  
QY 105 SPQPTIIVRGTESSVRFVNSGENTSPNSVHLHGSFRRAPDDGAEEDTQGEKDYKYY 164  
D 126 spqptifgekgresvrvfnka--twesslhlngsrtapwgaevltvnpgeykdyyp 183  
XX  
QY 165 NROAARMLWHDHAMSITTAENAYMGAGVYMIODPADALNPSGYGEFDIPLVLTAKRY 224  
D 184 naqagrfswyhdhametaenaayngaggyllhndpaedslglsygyeydiprlllskgy 243  
XX  
QY 225 MADGTLESTNGEVSFWGDYIQVNGOPWMLNQPCKRYRREFLNAVSRSAVLALATSED 284  
D 244 nsdgtlftskgetqslwgdvlyngvwpvfyfdeprkkyrriflidaavsrsfslyfvdad 303  
XX  
QY 285 SEFRLPFOVIAADGGLLEGVDVDTLYISMAERWEVVIDSTFAGOSIDIRNLPGADG 344  
D 304 edrlpftqvlaasdgllleevtkslvlsaeylelldsdfegktelempavagyl 363  
XX  
QY 345 VEPEPDNTDKVMRF-VVDEVLESPTSEVPANLRDVPFPEGGMNDPNDTDFTFTFGRA 403  
D 364 lervnyddctkvmrfvnaegslspdsctsvpstrldvrfps-----tsltldsfart 418  
XX  
QY 404 NGQWTINGVTFSDVENRLRNVRPRTVEIWRLENNNGMTWHPVHILVDFRLSR----- 458



```

|||||
419 agwgsingvfdsventrllanvplgtvqlwqlinaagwphlphlhdvfkilstrgaac 478
459 -SPARGVEPEAAGLKDVVWLARREVVYVAHAAPFGVYMLCHNLIHEDHMMAAFNV 517;
479 galtrvgvpepsaglkdvvylykgelvvvvaahapwgyvmtfchnlthednmmaafnv 538
518 TVLGDYGYNTFEIDPMEPLMRPRPFLLGFEENGSGDFSELATDRIQEMASFNYPAAQD 577
539 tvlpdygynstaldpmdqdefrakrakyvndvrenafstdeleaqvqlmasyljpydnpd 598

RESULT 6
AA95538
ID AA95538 standard; Protein: 627 AA.
XX
XX AA95538;
XX
XX 10-OCT-2000 (first entry)
XX
XX Bipolaris spicifera phenol oxidizing enzyme.
XX
XX Phenol oxidizing enzyme; detergent; bleaching.
XX
XX Bipolaris spicifera.
XX
XX WO200039306-A2.
XX
XX 06-JUL-2000.
XX
XX 20-DEC-1999; 99WO-EP10287.
XX
XX 23-DEC-1998; 98US-0220871.
XX
XX 23-JUN-1999; 99US-0338723.
XX
XX (UNIL ) UNILEVER NV.
XX
XX (UNIL ) UNILEVER PLC.
XX
XX (HIND-) HINDUSTAN LEYER LTD.
XX
XX
XX Bodie EA, Van Der Velden S, De Vries CH, Wang H;
XX
XX WPI; 2000-514528/46.
XX
XX N-PSDB; AAA50020.
XX
XX
XX Detergent composition comprising novel phenol oxidizing enzyme obtained
XX
XX from fungus or bacteria, useful for pulp and paper bleaching, bleaching
XX
XX color of stains on fabric and for anti-dye redeposition
XX
XX
XX Claim 8; Fig 3; 45pp; English.
XX
XX
XX The present sequence is that of the Bipolaris spicifera
XX
XX phenol oxidizing enzyme. The invention relates to detergent
XX
XX compositions comprising novel phenol oxidizing enzymes that have at
XX
XX least 60% identity with the phenol oxidizing enzyme of Stachybotrys
XX
XX chartarum (see AA95537), and which are obtained from a bacterium,
XX
XX yeast or non-Stachybotrys fungus, especially B. spicifera,
XX
XX Curvularia pallescens (see AA95539) and Amerosporium atrum (see
XX
XX AA95540). The phenol oxidizing enzyme is capable of modifying the
XX
XX colour associated with dyes or coloured compounds, and can be used
XX
XX for pulp and paper bleaching, for bleaching the colour of stains on
XX
XX fabric and for anti-dye transfer in detergent and textile
XX
XX applications. It may also be capable of modifying the colour in
XX
XX the absence or presence of an enhancer. Expression vectors and
XX
XX host cells comprising a nucleic acid encoding a phenol oxidizing
XX
XX enzyme, methods for producing the phenol oxidizing enzyme, and
XX
XX methods for constructing expression hosts are provided.
XX
XX
XX Sequence 627 AA:

```

Query Match 56.6%; Score 1762; DB 21; Length 627;  
 Best Local Similarity 61.6%; Pred. No. 1.7e-151;  
 Matches 329; Conservative 71; Mismatches 118; Indels 16; Gaps 7;

```

47 PLALVEPLIPPLKAPNTVPNPTGEDILYEMERFSHQIYDPLEPANMYGDGMS 106
67 pl-1frgpliprpkapknkltnpvtknkeiygelvklpfqgyrpslrparlygdygisp 125
107 GPPTIYPRGESVYRVFVNSCENTSPNSVHLGFSFRAFDGMAEDTTOGEYKDYVYV 166
126 gptliivprgeavvrfingdres--shlhgsprsfcdgwdmkmkgedydyppnn 183
167 QAARMLMYHHAASITRNNVYMGAGVYMTQDAEDLNLPSGIGEDTILVITAKRYNA 226
184 qaarflwyhhamvtaenaylfgagaylltdpaedalglpsgykydldlvlskyyna 243
227 DGTLESTNGEVSSEGVNIOVNGOPWMLNVOPRKYRFRFLNAVSSFLVATSEDS 286
244 dgltkesvgedksvwdgdlhvnngpwpflvneprkyrlfrlnaavsnfilyfvkqdn 303
287 TRLPFOYIAADGGLLEGVPVDTPLIYISMAERWEVVDSTFGAQSIDIRNLPGADGL 346
304 trlpfgyiasdagllthpvgtsdmyvaaeeyelvfdfayagqlldlrnfakangid 363
347 PERDNDKVMRFYVDEVLESPTSEVPANLRDVPFPEGGMWDANPTD-DETTFGANG 405
364 ddyantdkvmrfvvs-qvvdnsvpeqlsqif-----padktdidhtrfhftng 415
406 QWTINGVTFSDVENRLRNVPRTVEIWRLENNNGWTHPVHILVDFYLSR---STAR 462
416 ewrlnglfgfadvendrlyakvprglvewlenssgysphlrvhldvfrvaygdegr 475
463 GVEPEAAGLKDVVWLARREVVYVAHAAPFGVYMLCHNLIHEDHMMAAFNVTVLGD 522
476 gvmpyeaaglkdvvwlgrhctviveahyapwdgymfchcnlthbedqmmaatdvtklgn 535
523 XGYN-VTEFTDPMELMRPRPFLLGFEENGSGDFSELATDRIQEMASFNYPAAQ 575
536 fgyntctdlnpdpwrsarpfagdlrtarsgiffeslrayvnelaleqpyse 589

```

```

RESULT 7
AA96762
ID AA96762 standard; Protein: 627 AA.
XX
XX AA96762;
XX
XX 09-OCT-2000 (first entry)
XX
XX Bipolaris spicifera phenol oxidizing enzyme.
XX
XX Phenol oxidizing enzyme; colour; dye; modification; detergent; stain;
XX
XX pulp; paper bleaching.
XX
XX Bipolaris spicifera.
XX
XX WO200037654-A2.
XX
XX 29-JUN-2000.
XX
XX 20-DEC-1999; 99WO-US31009.
XX
XX 23-DEC-1998; 98US-0220871.
XX
XX 23-JUN-1999; 99US-0338723.
XX
XX (GENV ) GENENCOR INT INC.
XX
XX Wang H, Bodie EA;
XX
XX WPI; 2000-452191/39.
XX
XX N-PSDB; AAA51315.
XX
XX
XX New phenol oxidizing enzyme for modifying colors associated with dyes
XX
XX or colored compounds, is obtained from fungus and is encoded by a
XX
XX nucleic acid comprising a specific nucleotide sequence

```







```

XX Stachybotrys chartarum phenol oxidising enzyme.
DE Stachybotrys chartarum phenol oxidising enzyme.
XX
XX Phenol oxidising enzyme; detergent; bleaching.
KM
XX Stachybotrys chartarum.
OS
XX
XX WO200039306-A2.
PN
XX
XX 06-JUL-2000.
PD
XX
XX 20-DEC-1999; 99MO-EPI0287.
PF
XX
XX 23-DEC-1998; 98US-0220871.
PR
XX
XX 23-JUN-1999; 99US-0338723.
XX
XX (UNIL ) UNILEVER NV.
PA (UNIL ) UNILEVER PLC.
XX (HIND ) HINDUSTAN LEVER LTD.
XX
XX Bodie EA, Van Der Velden S, De Vries CH, Wang H;
PI
XX
XX WPI: 2000-514528/46.
DR
XX
XX N-PSDB: AAA50018, AAA50019.
XX
XX Detergent composition comprising novel phenol oxidising enzyme obtained
PT from fungus or bacteria, useful for pulp and paper bleaching, bleaching
XX color of stains on fabric and for anti-dye redeposition
XX
XX
XX Claim 2; Fig 5A-B; 45pp: English.
PS
XX
XX The present sequence is that of a phenol oxidising enzyme of
CC Stachybotrys chartarum MUC1 38898. A claimed detergent composition
CC comprises a phenol oxidising enzyme having at least 60% identity to
CC the present sequence, and preferably obtained from a bacterium,
CC yeast or non-Stachybotrys fungus, especially Bipolaris spicifera,
CC Curvularia pallescens or Amersporium atrum (see AAY9538-40). The
CC phenol oxidising enzyme is capable of modifying the colour
CC associated with dyes or coloured compounds, and can be used for
CC pulp and paper bleaching, for bleaching the colour of stains on
CC fabric and for anti-dye transfer in detergent and textile
CC applications. It may also be capable of modifying the colour in
CC the absence or presence of an enhancer. Expression vectors and
CC host cells comprising a nucleic acid encoding a phenol oxidising
CC enzyme, methods for producing the phenol oxidising enzyme, and
CC methods for constructing expression hosts are provided.
XX
XX
XX Sequence 594 AA:
SO

```

QY	343	LGVPPEPDNDKVAKEVYVDE-VLESDDTSEVPANLEADVFP--	EGGMWMDPANPTDDTET	398
Db	355	vgdgedearllevmrtfivssgtve---dsqdvpsltldvfpfhksg--- <td>pa----dkhf</td> <td>404</td>	pa----dkhf	404
QY	399	TFGRANGQMTINGVTSFSDVENRLLRNVRPRVDEIMRLNNNSGMGTHPYHILHYFVRYSR		458
Db	405	kfersnghyllndvgfadvnevrllakpeltgtvewelenssgswbphvhlhvdtkllkr		464
QY	459	STAGG-VEPTEAGAGKRYVYLARREVVYVEAHAPPGYYMLHCHNLHEDHDMAAATNV		517
Db	465	tgrgqvmypesaglkdvwlvgfgetclieahygpwtgtaymwhchllheddmavtnv		524
QY	518	TVLDDYGVNTEFTIDPEPMLRRPRFLTGFEGSGDFSELAITPRIDEMASFNFPYAQAD		577
Db	525	tameekylqgedfedpmpkrtavpyrnrdfharaagnfaesitarvaelaegpyrild		584
QY	578	D 578		
Db	585	e 585		
RESULT	13			
ID	AA96761			
XX	AA96761	standard; Protein; 594 AA.		
AC	AA96761:			
XX	09-OCT-2000	(first entry)		
DT	Stachybotrys chartarum	phenol oxidizing enzyme.		
DE	Phenol oxidizing enzyme; colour; dye; modification; detergent; stain;			
XX	pulp; paper bleaching.			
KW	Stachybotrys chartarum.			
XX	Mo200037654-A2.			
PN	29-JUN-2000.			
PD	20-DEC-1999;	99MO-US31009.		
PF	23-DEC-1998;	98US-0220871.		
PR	23-JUN-1999;	99US-0338723.		
XX	(GENMV ) GENENCOR INT INC.			
PA	Wang H, Bodie EA;			
XX	WPI: 2000-452191/39.			
DR	N-PSDB; AAA51313, AAA51314.			
DR	New phenol oxidizing enzyme for modifying colors associated with dyes			
XX	or colored compounds, is obtained from fungus and is encoded by a			
FT	nucleic acid comprising a specific nucleotide sequence			
XX	Claim 1; Fig 5A-B; 45pp; English.			
PS	This is the sequence of Stachybotrys chartarum phenol oxidizing enzyme.			
XX	Phenol oxidizing enzymes encoded by nucleic acid sequences which			
CC	hybridize to the coding DNA are claimed, as long as the enzyme is capable			
CC	of modifying the colour associated with dyes or coloured compounds. The			
CC	enzymes are useful in detergent compositions and for modifying colors			
CC	associated with dyes or colored compounds which occur in stains in a			
CC	sample. The enzymes are also useful for pulp and paper bleaching,			
CC	anti-dye transfer in detergent and other textile applications.			
XX	Sequence 594 AA:			
Query Match	55.1%;	Score 1716.5;	DB 21;	Length 594;
Best Local Similarity	61.0%;	Pred No. 2.2e-147;		



OY 496 VYMLCHNLHEDHMAAFNV 517  
 1: ||| | | | | |  
 Db 451 qvyhchylehsslgmaqlv 472

## RESULT 15

AAG90799  
 ID AAG90799 standard; protein; 511 AA.

AC AAG90799;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 4553.

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
 organic acid synthesis.

OS Corynebacterium glutamicum.

PN EP1108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

PA (KYOW ) KYOMA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H,  
 Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR MPI; 2001-376931/40.

DR N-PSDB; AAH66018.

PT Novel polynucleotides derived from Coryneform bacteria, for identifying  
 mutation point of a gene, measuring expression of a gene, analysing  
 expression profile or pattern of a gene and identifying homologous gene

PS Claim 17; SEQ ID NO: 4553; 246pp + Sequence Listing; English.

CC The present invention provides a number of nucleotide and protein sequences from the Coryneform bacterium Corynebacterium glutamicum; these are useful for identifying the mutation point of a gene derived from a mutant of coryneform bacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from Coryneform bacterium, and identifying a homologue of a gene derived from Coryneform bacterium. Coryneform bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a protein described in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office.

XX Sequence 511 AA;

Query Match 16.2%; Score 504; DB 22; Length 511;

Best Local Similarity 29.5%; Pred. No. 5.9e-37; Mismatches 232; Indels 72; Gaps 18;

OY 7 GAVALGLAVIGG-----SSVDARSVAGRSTDPGSLTKROTOLSPPLALYEVPLPIPL 60

Db 12 gagvlatvgaqyivacssdvrygggpr-----tlpippa 49

OY 61 KAPTVVNPNTGEDILYEMKIRFESHQIYPDLEPANMGVYDGKSPGPIIVPGTESVY 120

Db 50 dl-----gtregsav-hfaiaeqtesqllpdytklw-gfngtlhlpvlvkkgdvhw 102

OY 121 RFVNSGENTSPNSVHILHSGFSRAFPDGMADDTQPGEXKRYYPNRGAAMLYHDIAMS 180

Db 103 dvinldemt-lvhwghmklpalaaggphspilpgqtwspwlvandaatlwyhphthg 160

OY 181 ITAENAYMGAGVYMIIDPDAEDALNLPSCGGEFIPVLTKAKRYNADGTLFTSTNGEVSSF 240

Db 161 ltgthayrtaagmliavedatdkldipreyvddipvlmhthfledgsidedlpolgl 220

OY 241 WGDVIQVNGQPWMLNVQPRKRRFRFLNAAVSRSFALYLATSEDESTRLPFOYIADGGL 300

Db 221 lqdtptanglnahfdatctrrvfrvngsmrfynlaf-----sdlr-lfyiasdsgl 274

OY 301 LEGVVDITLYISMAEWEVVIDESTFAGSIDRNLPAGDGLVE-----PEPDNTDKV 355

Db 275 ldepqdrtllalpggetweivele--pgeadvlesvgfednygvpddefvptdgmddf 332

OY 356 MRFVDEVLSPPTSEVPANLRDVPPECGNMPDPANPTDDETFEGRANGQWTRNGVTF 415

Db 333 qltlitg-psdaaagpa-----lpgylvkstep-dvidatertf-----lmttsin 378

OY 416 DVENRLLR---NWPRDTVEIWRLENNNSGWTHPVHILVDFRVLRSRTANGVPEYEAGL 472

Db 379 dlgmdmqvrdvldhdqpevwlvclndsdwphmfhvndarfkvl-kfegtdvelfn-dgw 436

OY 473 KDYVWL---ARREVVYEAHAYRPPGYYMLCHNLHEDHMAAFNVYGLD 522

Db 437 kdtvqlppgatatlavelfgnyppqpwpymyhcimlyhedgmmgqfivvepgd 489

Search completed: October 3, 2002, 10:56:39  
 Job time: 143 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:16 ; Search time 34.9 seconds  
(without alignments)  
1605.160 Million cell updates/sec

Title: US-09-656-640A-2

Perfect score: 3114

Sequence: 1 MISQAIGAVLGLAVIGSS.....IQEMASFNPRYAQADDAE 583

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71:1\*  
1: PIR1:1\*  
2: PIR2:1\*  
3: PIR3:1\*  
4: PIR4:1\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2010	64.5	572	2 B48521	bilirubin oxidase
2	731.5	23.5	513	2 F69604	spore coat protein
3	489.5	15.7	568	2 B86364	hypothetical prote
4	460	14.8	533	2 AC0414	probable exported
5	442	14.2	527	2 C70397	periplasmic cell d
6	431	13.8	591	2 G96734	spore coat protein
7	416.5	13.4	494	2 AC3582	probable blue-copp
8	413.5	13.3	516	2 G90644	hypothetical prote
9	413.5	13.3	516	2 G85495	hypothetical prote
10	412.5	13.2	516	2 C64735	probable copper-bl
11	402.5	12.9	536	2 AF0023	probable copper-bl
12	339.5	10.9	474	2 AD0083	probable cell divi
13	322	10.3	470	2 E91116	suppressor of fsti
14	322	10.3	470	2 E85961	suppressor of fsti
15	319	10.2	470	2 G65088	sulfu protein precu
16	310	10.0	470	2 AD0888	Sulfu protein (limp
17	288.5	9.3	1662	2 T18540	moifa protein precu
18	287.5	9.2	513	2 G81298	probable periplasm
19	259.5	8.3	463	2 G83175	probable metallo-o
20	256	8.2	500	2 B83910	hypothetical prote
21	247.5	7.9	622	2 S56214	probable membrane
22	247	7.9	311	2 H64157	sulfu protein homol
23	246	7.9	504	2 F70813	hypothetical prote
24	218.5	7.0	721	2 H82528	L-ascorbate oxidas
25	199.5	6.4	551	2 T02752	probable laccase (
26	199	6.4	529	2 S68120	laccase (EC 1.10.3
27	196.5	6.3	622	2 S62580	probable multi-copp
28	196.5	6.3	632	2 F83387	copper resistance
29	193.5	6.2	635	2 A36868	coppa homolog - Xan

30	193	6.2	529	2 S49120	laccase (EC 1.10.3
31	192	6.2	619	1 KSNCL0	laccase (EC 1.10.3
32	191	6.1	580	2 F84828	probable laccase (
33	190.5	6.1	557	2 JC5229	laccase (EC 1.10.3
34	190	6.1	636	2 A55428	ferrioxidase precu
35	188.5	6.1	533	2 S62371	laccase (EC 1.10.3
36	188.5	6.1	570	2 H87368	copper-binding pro
37	188	6.0	548	2 S18746	laccase (EC 1.10.3
38	188	6.0	572	2 S68119	laccase (EC 1.10.3
39	187.5	6.0	527	2 JC5357	laccase (EC 1.10.3
40	187.5	6.0	621	2 S72493	laccase (EC 1.10.3
41	186.5	6.0	520	2 S59533	laccase (EC 1.10.3
42	183.5	5.9	520	2 JC5356	laccase (EC 1.10.3
43	183	5.9	619	1 KSNCLT	laccase (EC 1.10.3
44	182.5	5.9	587	1 KSKV40	L-ascorbate oxidas
45	181	5.8	554	2 T01240	laccase (EC 1.10.3

## ALIGNMENTS

## RESULT 1

B48521  
bilirubin oxidase (EC 1.3.3.5) - fungus (Myrothecium verrucaria)

C:Species: Myrothecium verrucaria

C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000

C:Accession: B48521; A48521

R:Koike, S.; Ando, K.; Kaji, H.; Inoue, T.; Murao, S.; Takeuchi, K.; Samejima, T.

J. Biol. Chem. 268, 18801-18809, 1993

A:Title: Molecular cloning of the gene for bilirubin oxidase from Myrothecium verruca

A:Reference number: A48521; MUID:93366794

A:Accession: B48521

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-572 <KOI1>

A:Cross-references: GB:D14081; NID:q436236; PIDN:BA03166.1; PID:q456710

A:Note: Sequence extracted from NCBI backbone (NCBIN:136730, NCBIIP:136732)

A:Accession: A48521

A>Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-572 <KOI2>

A:Cross-references: GB:D14081; NID:q436236; PIDN:BA03166.1; PID:q456710

A:Note: Sequence extracted from NCBI backbone (NCBIN:136728, NCBIIP:136729)

C:Keywords: oxidoreductase

Query Match	Score	2010	DB 2	Length	572
Best Local Similarity	65.4%	Pred. No. 1	1e-135		
Matches	383	Conservative	71	Mismatches	110
				Indels	22
				Gaps	10
QY	1	MISQAIGAVLGLAVIGSSVDARSVAGRTDMSSG-LTRKOTOLSPPLALYEPPLIPP	59		
DB	1	MEKHTLCAALSLT-LFNSNAVAQASVPV-ETSPATGHLFRVAQISQYPPFVPLIPP	57		
QY	60	LKAPN-TVPNPTNGEDILLYEMEIRPFSHOIYPLLEPANNMGYGMSPGFTIYPRCTES	118		
DB	58	VKQRLTVNPNVNGQELMYEVEIKPETHOYVPLDGLADLVGDMGSPGFTIYPRCTES	117		
QY	119	VRFVNSGENTSPNSVHLHGSFSRAPPDGMNAEDTTPGEXKDYPPNROAARMLYDHA	178		
DB	118	VVRFINNAE--APNSVHLHGSFSRAAPDGMNAEDTTPGEXKDYPPNROAARMLYDHA	175		
QY	179	MSITAEAAVYMGQGVVITODPAEDALMLPSGCGFEDIPVLVIAKRYADGLTSTNGEVS	238		
DB	176	MHTTAEAAVYMGQGVVITODPAEDALMLPSGCGFEDIPVLVIAKRYADGLTSTNGEVS	235		
QY	239	SFMGDVYQVNGQPPMPLNVOPRKRYRFRFLNAAVRSFALVTSSEDETRLPVQVIAAG	298		
DB	236	SFMGDVYHVNQGPMPFKNVPRKRYRFRFLNAAVRSFALVTSSEDETRLPVQVIAAG	295		
QY	299	GLLEGVDTDTLYISMAERREVVIFSTFAGQSIDIRNLPGA-DGLGVEPEFDNTDKVMR	357		
DB	296	GLLEHPTDTSILYISMAERREVVIFSTFAGQSIDIRNLPGA-DGLGVEPEFDNTDKVMR	355		



```

Query Match Similarity 14.8% Score 460; DB: 2; Length 533;
Best Local Similarity 28.3%; Pred. No. 56-25; Mismatches 163; Conservative 62; Indels 140; Gaps 22;
Matches 163; Conservative 62; Indels 140; Gaps 22;

QY 28 GASTMPGSLIKRQQLSPALLAYEPLPIPLKAPNTVNPNTGDDILYEIMEIRFSH 87
Db 14 GATSLP-LMSRAA-----LADFSLPIPLLPDA-----KG 46
QY 88 QIYPLPEANNV-----GYDGMSPGPTIIVPRGTSYRVFNGSGENTSPNSVHLH 137
Db 47 KNLNLIGTGSYVWLSESTATQWTGNGNLLGPAIRLORGAQVNTDITNALPEAT--TVHWH 104
QY 138 GSFSPAPFGMAEDFTQPEYKDYVYPNQQAARMLYHHHNASITFENAMVMOAGVYMQ 197
Db 105 GLEIRGEVNDGCGQALIDPEAKRQYTFVAYEQPATCGFHHTHSKITHQVAMKGLGLVLID 164
QY 198 DPAEDALNPSYSGEDFDIYLTAIRYNADGTLFTNGEVS-----SFWGVYIOV 247
Db 198 DPAEDALNPSYSGEDFDIYLTAIRYNADGTLFTNGEVS-----SFWGVYIOV 247

```

Db 165 DSDSEFLPLPKQMGVDDIIPVLIQDK-----LDDKHQVQYQYQDVMYTAAGWFGDRLNT 217

Qy 248 NCQPPMPLNVOPRKI-KREFLNAAVSRSFALYLATSEDSFRLPFOVIADAGLLEGVVD 306

Db 218 NCVPYPO-QITRPQVWRRLRLNGCARSILNALSIDGR-----PMYVIAASDGLLAEVSV 270

Qy 307 TDTLYISMAEREEVVIDSTFAGOSIDIRNLPGAD-GIGVEPEPNDIKVWF-----VV 360

Db 271 VELLPLNGEREVLVD--TRDQSLDLVTLPTVMQMTLAP-FDQPLPLRLIQPSLAIG 327

Qy 361 DEVLESPDSEVPANILRDPV-----FP 382

Db 328 SQVL--PESLVYIPELADVTGYQERWFOPLMDPKIDMLGMALVARYGMKAKMAGNMNHG 385

Qy 383 EGGNNDPANPFD-----DEFTTEGRANGQMTINGVFSVDENRLRNVP 426

Db 386 DMGADHDGHRPDSOGKKMGDHTMGAPAFNFSHAN--RINGKAFSMTPEAF--DAK 440

Qy 427 RDTVEIWRLENNSCWMTPVHILVDPRVLRSRARGVEPYE-AAGKADVWL--ARREV 483

Db 441 QCKYRKWTISGEGDMLHPFHVHGTQFRIL--TENGRPPAEHRGRKWDIVVEGARSEI 497

Qy 484 VYVEAHYAPFPGVYMLHCNLIHEDHDKMAAFNVY 518

Db 498 LVRFNVYLAAPSPPYMAHCHLLEHEDTGMLGFTYS 532

RESULT 5

C70397

periplasmic cell division protein (SufI) - Aquifex aeolicus

G.Species: Aquifex aeolicus

C.Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999

C.Accession: C70397

R.Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V. Nature 392, 353-358, 1998

A.Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A.Reference number: A70300; MUID:98196666

A.Accession: C70397

A.Status: preliminary; nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-527 <AOP>

A.Cross-references: GB:AE000724; NID:g2983585; PIDN:AA07157.1; PID:g2983586; GB:AE000724

A.Experimental source: Strain VFS

C.Genetics:

A.Gene: sufi

Query Match 14.2%; Score 442; DB 2; Length 527;

Best Local Similarity 28.7%; Pred. No. 9,4e-24;

Matches 165; Conservative 77; Mismatches 204; Indels 128; Gaps 29;

Qy 10 ALGLAVIGGS--SYDARSVAGRSTDMPSGLRKROTQLSPRLALYEVLPPLKAPN--- 64

Db 14 ALGFSVGGLSLSCGGGTTSSSGGSGTSLKSKSLNIPGYLF-----PDGQR 62

Qy 65 -----TVPNPNGEDILYEMEIRPFHSQIYFDLEPANMVGYDGMSPGPTIIVP 113

Db 63 VSIATKWTITLLEIPEKSF--DMLYEID-----NEYNPVFLR 98

Qy 114 RGTESVVRVY-NSGENTSPNSVHLHGSFRAFP--DCMAEDTTQPGE---YKDYVYNNRQ 167

Db 99 KQGFPSADFEVNNSGED--SIITHHG--FRAPMKSDGHPYAAVAGDEVSYSDPTIIDS 153

Qy 168 AARMMLYHDHAMSTIAENAYMGAGVYMIQDPADALN--LPSTGGEFDIPLVLTARYN 225

Db 154 GT--YFHPHGHGRGTYGYGLAGMIIEDEDENKQALDLEYGVIDIPLIQTDTFD 211

Qy 226 ADGTL-FSTNGEVSFWGDVYIQVNGQPPMLNVOPRKRYRFFLNAAVSRSFALYLATSED 284

Db 212 SSGQVLYNPMKMG-GFMQDITLVNLTLPNPNYDVERKTIYRFFILNGSNAPRYRLALRGNQ 270

Qy 285 SETRLPFOVIADAGLLEGVDDTLYISMAEREEVIDF-STFAGOSIDIRNLPG--- 339







```

QY 250 QPWLPLNQPCKRYRFRFLNAAVSRSEFALYLTSEDESETRLEPOVIAAAGGLLEGVPYDIT 309
Db 218 AQPPEVEYSRGMWRRLRLNLSAARYYTLQSLSDGR-----PLYVASDQGLPAPVAVOQ 271
QY 310 LYISMAERWEVYVIDF-----STFAGQSIDI-----RNLPGADGL-----GVEP 347
Db 272 LSLAAGEERREVVYIDMSQNEVSIITFAGEESAGIMDRLRGLFEPSSILITVLTLKPTGLP 331
QY 348 EEDNDTKX-MRFVYDEVLESPTSEVPANLRDVPFPEGGNMDPAFPTDEFTFETGRANQ 406
Db 332 LV--TDNLPMLRLADQIIIEG-----SVIRSREQLQDNLNPGIN-----GAI 370
QY 407 WTINGVFESDVENRLLRLRVPRDJVEIMRLENNNSNGTAPVHIILHVDFFRLSRSTARVEP 466
Db 371 WDMNRY--DYQ-----AQOCTMERWITIHAD--MPQAFHILQGVSLYKSVNGAAMAA- 417
QY 467 YEAGLRKRVVWL 478
Db 418 -EDRGWKDJVAV 428

```

## RESULT 13

suppressor of ftsI [imported] - *Escherichia coli* (strain O157:H7, substrain R1MD 0509952)  
 C:Species: *Escherichia coli*  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 C:Accession: E91116  
 R:Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.-G.,  
 Gasawara, N., Yasunaga, T., Kuhnara, S., Shiba, T., Hattori, M., Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic  
 A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: E91116  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-470 <HAV>  
 A:Cross-references: GB:BA000007; PIDN:BA37324.1; PID:G13363373; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain R1MD 0509952  
 C:Genetics:  
 C:Gene: ECs3901

Query Match	10.3%;	Score 322;	DB 2;	Length 470;
Best Local Similarity	25.3%;	Pred. No. 2.9e-15;		
Matches 131;	Conservative 68;	Mismatches 204;	Indels 114;	Gaps 21;

```

QY 10 ALGAVLSSGVDRASVAGRTDPSGLTKRQQLSPDLALVEYPLIETPLKAPNTVPNP 69
Db 11 ASGIALACAGAPLKASNAGO-----QQLPVPPLL-----ES 42
QY 70 NTGEDILYEMEIRPESHQIYPDLEPANWGYDMSGPITIIVRGTESVVRVNSGENT 129
Db 43 RRGQPLF---MTYOR-AHWSTPGSTRASVINGINRYLGPFIWKKGD--VKLTISNRLT 96
QY 130 SPNSVHLHGSGSRAPDFGMAEDTTQPEBYNDYYPNQOARMLYHDMHASTIAENNYMG 189
Db 97 ENVSMTYAVGLOVPEPLMGGRPMRMSPADNAPVLIQONATMTYANTENRTPAQOYYNG 156
QY 190 QAGYMTIODPAEDALNLPSCYGEEDILVLTAKKRYNADGTLFTENGSEVSFWGDVAVNNS 249
Db 157 LAGMNLVEDEYSKSLPIPNHGYVDPEVITIQDKRLDNFGRPENNEPSGSGFVGDTLLVNS 216
QY 250 QPWFMLVQPPKKYFRFLNNAVNSFSLYLTATSDSETRLPFOVIADGGLLEGVDYDT 309
Db 217 VQSYVEVSRCQWBLRLINMSNSRRYQLQMSDGR-----PLHYISDQGFLLPAPVSKQ 270
QY 310 LYIEMAEWEVVIDFST-----FAGOSIDIRNL-----PGADGL 343
Db 271 LSLAPGERRELIVMSKNGDEVSTICGEAASLYDRIRGFEPFSSILVSTVLTILRP-----T 326
QY 344 GVPEDFNTDKV-MRFVDEVLESDPTSEVPANLRDVPFEGCGMNDPANTDDETFYFGR 4020

```

D6	327 GLELLV--TSLPRLRPPTLIMAGS-----PIRSRDLSLD-----DPG-----	363
OY	403 ANGO-WTINGTSFSDVENRILLRNPRDVEIWRLENNSNGWTHPHHLHYDFVLVSRTA	461
	:   :   :	
D6	364 INGLMDVNRI---DV-----TAQOSTBERWIYRADE---PAFHIEGMPOI---RWV	408
OY	462 RGVETPY-EAAGLKRYVM---ARREVVYYVAHIAAPP	494
	:   :   :   :   :	
D6	409 NGAMPFPEDRGWKDTAVWDGVELLYEVGGPSNAHP	445

## RESULT 14

suppressor of tsr1 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)  
 C.Species: *Escherichia coli*  
 C.Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C.Accession: E85961  
 R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May, M.E.;  
 Miller, L.; Gotohck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.  
 Nature 409, 529-533, 2001  
 A.Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A.Reference number: A85480; MUID:21074935; PMID:11206551  
 A.Accession: E85961  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-470 <STO>  
 A.Cross-references: GB:EM005174; NID:G12517586; PIDN:AAG58153.1; GSPDB:GN00145; UWGP:  
 A.Experimental source: strain O157:H7, substrain EDL933  
 C.Genetics:  
 I.Gene: stx1

Query Match	Score	DB 2	Length
10.38;	322;		470;

Best Local Similarity 25.3%; Pred No. 2.9e-15;  
Matches 131; Conservative 68; Mismatches 204; Indels 114; Gaps 21

```

11 ASGIALCAGANPLKASAGQ-----QQPLEVPPLL-----ES 42

```

09 70 NTGEDILYEMETRPFHQIYPDLEPANWGVGDMSPPTIIVPRGTESVVRVNSGENT 12

Dh 43 RRCODL F---MTTVOR-AHWSFTPGT RASVWG TNGRYLGPT RVWKGDD--VKI IYSNRI T 96

[illegible]

QY 130 SENSVALNOST SKAFFU GWAEI I QFGEI INDI I PRNKQAKMLMI HDHAMSITAE NAIMG 18

| : : | | | | | | | : : | |

Db 97 ENVSMTVAGLQVPEPLMGGPARMSPNADNAPVLPIRONATLWYHANTPNTAQOYNG 15

QY 190 QAGVYMIQDPAEDALNLP SGYGEFDIPLVLTAKRYNADGTLFSTNGEVSSFWGVDVIQVNG 24

Db 157 LAGMWLVEDEVSLSLPINHYGVDDFPVILIQDKRLDNFGTPEYNEPGSGGIVGDTLLVNG 21

QY 250 QPMPMLNVQPRKYRFRFLNNAVSRSFALYLATSEDESETRLPFQVIADGGLLEGPVDTDT 300

```
Db      217 VQSPYVEVSRGWRLRLINASNSRRYQLWMSDGR-----PLHVISGDGELLPAVSVKQ 27
```

0x 310 IYISMAFFWEWIDST-----FAGOSIDIBNT-----PGADT 34

[illegible]

DD 2/1 USUNFGENEIDIVDM3NGDEVSICGEMASVDRKINCFEFS3ILV3ILVLELNF---1 32

QY 344 GVEPEFDNTDKV-MRFVDEVLSPDTSVPANLKDVPPEEGGNWDPANPIDDEFTTFFGR 400

Db 327 GLLPV--TDSLPRLLPTEIMAGS-----PIRSRDISLGD---DPG----- 36

403 ANGQ-WTINGVPSDEVENRLLRNVPDQVEIQRLENNNGWTHPVHILYDFRVLSRSTA 46

```

Db      364   INGLMDVNR I--DV-----TAQQGTWERWYRADE---PQAFHIEGVMFQI---RNV 400
          |||..|||

```

QY 462 RGVEPY-EAAGLKDVWL---ARREVVYEAHYAPFP 494

db 409 NGAMPFEDRGWKDTVWVDGOVELLYFEGOPSWAHP 445

RESULT 15

G65088

sufi protein precursor - Escherichia coli

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999

C:Accession: G65088; S20461

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.: Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: G65088

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-470 <BLAT>

A:Cross-references: GB:AE000384; GB:U00096; NID:q2367186; PIDN:AC76053.1; PID:q1789394;

A:Experimental source: strain K-12, substrain MG1655

R:Coleman, J.

Mol. Gen. Genet. 232, 295-303, 1992

A:Title: Characterization of the Escherichia coli gene for 1-acyl-sn-glycerol-3-phosphat

A:Reference number: S20459; MUID:92212294

A:Accession: S20461

A:Molecule type: DNA

A:Residues: 1-89 <COL>

A:Cross-references: GB:M63491; NID:q147296; PIDN:AAA24398.1; PID:q147299

C:Genetics:

A:Gene: sufi

Query Match 10 28; Score 319; DB 2; Length 470;

Best Local Similarity 25.38; Pred. No. 4.8e-15;

Matches 131; Conservative 67; Mismatches 205; Indels 114; Gaps 21;

```

QY 10 ALGLAVIGSSVDARSVAGSTDMPSGLTRQQLSPPLALYEVPLPPLKAPNTVPNP 69
DB 11 ASGIALCAGAVPLKASAAQ-----QQPLVPPL-----ES 42
QY 70 NTGEDILYEMETIRPSSHQIYDLEPANVGYDGMSPGPTIIVRGTESVRFVNSGENT 129
DB 43 RRGQPLF--MTVQR-AHMSFTPGTRASVWGINGRYLGPTRVWKGD--VKLIYSNRLT 96
QY 130 SPNSVHLHGSFRAPFDGAEEDTQGEYKDYYPNRQAARMLWYHDHAMSTIAENAYMG 189
DB 97 ENVSMIVAGLVQVGPLMGSPARMSPNADWAPVLPTRQNAATLWYHANTPNRTAQQVYNG 156
QY 190 QAGVYMIQDPADALNLPSSGYEFDIPLVLTAKRYNADGLTFSTNGEVSFWGDVIOVNG 249
DB 157 LAGMMLVEDEVSKSLPIPMHYGVDDPFIYIQDKRLDNFGTPEYNERGSGGFVQDITLVNG 216
QY 250 QPWPMLNVQPRKRYRFRFLAAVSRSFALYLATSESESTRLPFOVIADAGGLEGPVDITP 309
DB 217 VQSPYVEVSRGWVRLRLNLSNSRRYLOMNDGR-----PLHVISGDGFLPAPVSVKQ 270
QY 310 LYISMAERWEVVIDFST-----FAGQSIDIRNL-----PGADGL 343
DB 271 LSLAPGERREIILVDMNGDEVISITGCEAASIYDRIRGFEPSIILVSTLVLTLP-----T 326
QY 344 GVEPEFDNTDKY-MRFVVDVELESPTSEVPANLRDVPPEGGMNDPANPTDDEFTFGR 402
DB 327 GLLPLV--TDSLPMRLLPFEIMAGS-----PIRSRDISLGD-----DPC-----*363
QY 403 ANGO-WTINGVTFSDEVENFLRNPRDIVEIMBLENNSGWTHPVHILVDFEVLRSSTA 461
DB 364 INGQLMDVNRVRI--DV-----TAQGTWERWTVRADE--POAFHIEGVMEFI--RNV 408
QY 462 RGEVPEY-EEAGLKDYYVWL--ARREVVYVEAHYAPRP 494
DB 409 NGAMPPEPDGRGKDVWVDGVYELLVYEGQPSMAHP 445

```





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:55:33 ; Search time 34.9 Seconds

(without alignments)  
1574.874 Million cell updates/sec

Title: US-09-656-640A-4

Perfect score: 3067

Sequence: 1 MFKHTLGAALSLFLFNSNAV.....AVTERIQTMAEYPPYAAAE 572

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: / 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3067	100.0	572	2 B48521	bilirubin oxidase
2	713	23.2	513	2 F69604	spore coat protein
3	506	16.5	568	2 B86364	hypothetical prote
4	470	15.3	591	2 G96734	spore coat protein
5	437.5	14.3	533	2 AC0414	probable exported
6	403	13.1	494	2 AC3582	probable blue-copp
7	401	13.1	516	2 C64735	probable copper-bi
8	400.5	13.1	527	2 C70397	periplasmic cell d
9	398	13.0	516	2 G90644	hypothetical prote
10	398	13.0	516	2 G85495	hypothetical prote
11	398	13.0	536	2 AF0523	probable multicop
12	333	10.9	474	2 AD0083	probable cell divi
13	318.5	10.4	470	2 E91116	suppressor of ftsi
14	318.5	10.4	470	2 E91116	suppressor of ftsi
15	315.5	10.3	470	2 G65088	su1 protein precu
16	311	10.1	470	2 AD0888	su1 protein (impo
17	276.5	9.0	1662	2 T18540	mo1a protein precu
18	276	9.0	513	2 G81298	probable periplasm
19	261	8.5	463	2 G83175	probable metallo-o
20	260	8.5	500	2 B83910	hypothetical prote
21	241.5	7.9	504	2 F70813	hypothetical prote
22	228	7.4	311	2 H64157	su1 protein homol
23	215	7.0	520	2 S59533	laccase (EC 1.10.3
24	212	6.9	520	2 J53536	laccase (EC 1.10.3
25	209	6.8	557	2 J53229	laccase (EC 1.10.3
26	207	6.7	527	2 J53357	laccase (EC 1.10.3
27	205.5	6.7	520	2 A35883	laccase (EC 1.10.3
28	205.5	6.7	520	2 B35883	lignolytic pheno
29	205	6.7	609	1 KSPSCY	copper resistance

30	201.5	6.6	580	2 F84828	probable laccase (
31	199.5	6.5	548	2 S18746	laccase (EC 1.10.3
32	198	6.5	551	2 T02752	probable laccase (
33	197.5	6.4	529	2 S68120	laccase (EC 1.10.3
34	197.5	6.4	554	2 T01240	laccase (EC 1.10.3
35	197	6.4	624	2 A36962	laccase (EC 1.10.3
36	195	6.4	721	2 H82528	L-ascorbate oxidas
37	194	6.3	619	1 KSNCL0	laccase (EC 1.10.3
38	193.5	6.3	622	2 S56214	laccase (EC 1.10.3
39	191	6.2	553	2 T45959	probable membrane
40	191	6.2	605	2 S52253	laccase-like prote
41	189.5	6.2	636	2 A55428	copper resistance
42	189	6.2	533	2 S62371	ferroxidase precu
43	187.5	6.1	621	2 S72493	laccase (EC 1.10.3
44	185.5	6.0	529	2 S49120	laccase (EC 1.10.3
45	184.5	6.0	611	2 C82845	laccase (EC 1.10.3
					copper resistance

#### ALIGNMENTS

Query Match	100.0%	Score 3067	DB 2	Length 572
Best local similarity	100.0%	Pred. No. 1.8e-217		
Matches 572	Conservative	0	Mismatches 0	Indels 0
Gaps 0				
1 MFKHTLGAALSLFLFNSNAVQASPVETSPATGHLKRVAAQISPOVPMFTVPIPIPVVQ 60				
1 MFKHTLGAALSLFLFNSNAVQASPVETSPATGHLKRVAAQISPOVPMFTVPIPIPVVQ 60				
61 PRFTVTPVNGOEIWEYVEIKPFTHOVYPDLGSADLVGDSGSPPTFOVPGVETVVR 120				
61 PRFTVTPVNGOEIWEYVEIKPFTHOVYPDLGSADLVGDSGSPPTFOVPGVETVVR 120				
61 PRFTVTPVNGOEIWEYVEIKPFTHOVYPDLGSADLVGDSGSPPTFOVPGVETVVR 120				
121 FINNAEPNSVHLHGSFRAAFDGMADITEPGSFKDYVYPNROSARTLWYHDHAMHITA 180				
121 FINNAEPNSVHLHGSFRAAFDGMADITEPGSFKDYVYPNROSARTLWYHDHAMHITA 180				
121 FINNAEPNSVHLHGSFRAAFDGMADITEPGSFKDYVYPNROSARTLWYHDHAMHITA 180				
181 ENAYRQAGLYMLTDPDAEDALNPSGVGEFDIPMLITSQVYANGMLVTTNGELNSFWGD 240				
181 ENAYRQAGLYMLTDPDAEDALNPSGVGEFDIPMLITSQVYANGMLVTTNGELNSFWGD 240				
181 ENAYRQAGLYMLTDPDAEDALNPSGVGEFDIPMLITSQVYANGMLVTTNGELNSFWGD 240				
241 VIVVNGQPMPEKFNVEPRKRYFRFLDAVSRSGLYFADPDADITRPLPFKYIASDSGLLEH 300				
241 VIVVNGQPMPEKFNVEPRKRYFRFLDAVSRSGLYFADPDADITRPLPFKYIASDSGLLEH 300				
241 VIVVNGQPMPEKFNVEPRKRYFRFLDAVSRSGLYFADPDADITRPLPFKYIASDSGLLEH 300				
301 PADTSLYISMAERYVVEFSDYAGKTIELRNLGSSIGIGTDTDTYDNTDKVMRFVAD 360				
301 PADTSLYISMAERYVVEFSDYAGKTIELRNLGSSIGIGTDTDTYDNTDKVMRFVAD 360				
301 PADTSLYISMAERYVVEFSDYAGKTIELRNLGSSIGIGTDTDTYDNTDKVMRFVAD 360				

Oy 361 DTTPDTSVPAANLRDVPSPPTNTPTROFRGRTGPTWTINGVAFADYQNLNLANVPVG 420  
 |||||  
 Db 361 DTTPDTSVPAANLRDVPSPPTNTPTROFRGRTGPTWTINGVAFADYQNLNLANVPVG 420  
 |||||  
 Oy 421 TVERMELINAGNGWTHPHIHLVDKYSRTSGNNARVMPYEGSLKDYVWLGRTETVVV 480  
 |||||  
 Db 421 TVERMELINAGNGWTHPHIHLVDKYSRTSGNNARVMPYEGSLKDYVWLGRTETVVV 480  
 |||||  
 Oy 481 EAHVAPFPGVYVHFCHNLTHEDHMAAFNATVLPDGYNAVYVDPDMELMOARPEYLG 540  
 |||||  
 Db 481 EAHVAPFPGVYVHFCHNLTHEDHMAAFNATVLPDGYNAVYVDPDMELMOARPEYLG 540  
 |||||  
 Oy 541 EFQASGQFVQAVTERIQTMAEYRPAAD 572  
 |||||  
 Db 541 EFQASGQFVQAVTERIQTMAEYRPAAD 572  
 |||||  
 RESULT 2  
 F69604  
 spore coat protein (outer) cotA - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
 C:Accession: F69604; A27393; S02538  
 R:Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloul, G.; Azevedo, V.; Berten  
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Erlington, J.; Fabret, C.; Ferrati, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Gall  
 lech, J.; Harwood, C.R.; Hensut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koether, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kutila, K.; Lapidus, A.; Lardinols,  
 A:Authors: Lauber, J.; Lavarovic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portelle  
 Rieger, M.; Rivalta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,  
 A:Authors: Schleich, S.; Schoeler, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Secor  
 akouchi, M.; Tamakoshi, A.; Tanaka, T.; Tempstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:9804033  
 A:Accession: F69604  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-513 <KUN>  
 A:Cross-references: GB:299107; GB:AL009126; NID:92632866; PIDN:CAB12449.1; PID:el182609;  
 A:Experimental source: strain 168  
 J.Donovan, W.; Zheng, L.; Sandman, K.; Losick, R.  
 J. Mol. Biol. 196, 1-10, 1987  
 A:Title: Genes encoding spore coat polypeptides from Bacillus subtilis.  
 A:Reference number: A27393; MUID:88011308  
 A:Accession: A27393  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-37 <DON>  
 R.Sandman, K.; Kroos, L.; Cutting, S.; Youngman, P.; Losick, R.  
 J. Mol. Biol. 200, 461-473, 1988  
 A:Title: Identification of the promoter for a spore coat protein gene in Bacillus subtil  
 A:Reference number: S02538; MUID:88286730  
 A:Accession: S02538  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-32 <SAN>  
 C:Comment: this spore coat protein is responsible for the characteristic brown pigment C  
 C:Genetics:  
 A:Gene: cotA; pig  
 C:Keywords: sporulation

Db 6 FVDALPIPTLKPVOO-----SKETTYEYMECTQHLDLPRTLMGNGLF 55  
 |||||  
 Oy 105 PGPTFOVPGVEYVAFINN-----AAFNVSYHLHSGSRAAF 142  
 |||||  
 Db 56 PGPTFOVPGVEYVAFINN-----AAFNVSYHLHSGSRAAF 142  
 |||||  
 Oy 143 DGMAE-----DITEPGS---KDYVPNRSQASRTLMYHDHAMHTAENAYRQOAGLIMLT 194  
 |||||  
 Db 116 DGYPEAMFSGDEQCGPRYKREYHYHPNQKALLMYHDHAMHTAENAYRQOAGLIMLT 175  
 |||||  
 Oy 195 DPAEDNLNPSGCGFERDMMILSKOYTANGNLVTTNGELN-----SEMGDYI 242  
 |||||  
 Db 176 DPEKRLKLPSS--DEYDVLTLTDRITNEDGSLFYSAPENSPSLPNSIYPAFCETI 233  
 |||||  
 Oy 243 HVGQMPFKNVPEPKRYRFRFLDAVSRSGFYFADTDAIDRLPCKVYASDGLLEHBA 302  
 |||||  
 Db 234 LVNGKWPYLEVPEPKRYRFRFLDAVSRSGFYFADTDAIDRLPCKVYASDGLLEHBA 287  
 |||||  
 Oy 303 DTSLYISMAERYEVVDESDYAGKTIELRNAGSISGIGTDTYDNTDKVRFVADDT 362  
 |||||  
 Db 288 KLSFSLAERDYDITIDTAVAGESIILANSAGCGDVNPETDAN-----IMQFRYTKPL 343  
 |||||  
 Oy 363 TQPDTSVPAANLRDVPSPPTNTPTROFR-----FGRTGPTWTINGVAFADYQNR 412  
 |||||  
 Db 344 AODDESRRKRYL--ASVPSQHERIONITFLKLAGTODEYGR--PVLLNNKRMADP--- 396  
 |||||  
 Oy 413 LLANVP-VGTVERMELINAGNGWTHPHIHLVDKYSRTSGNNAR----- 457  
 |||||  
 Db 397 -VTEPKVGTETLMSLINTPRG-THPHIHLVSVFRLDRPFIARYQSGELSYTGPAY 454  
 |||||  
 Oy 458 TWMPYSGKLDVWGLCRRETVVEAHYAPPGYVMPCHNLHEDHMAAFNAT 512  
 |||||  
 Db 455 PEPSEKGMKDTIOAHAGEVLRILAAATGPGYSGRYVWHCHLHEDYDMKRPMDIT 509  
 |||||  
 RESULT 3  
 B86364  
 hypothetical protein F19G10.5 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C:Accession: B86364  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
 ansen, N.F.; Hughes, B.; Hultzer, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzia  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: B86364  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-368 <STO>  
 A:Cross-references: GB:AE005172; NID:92462832; PIDN:AA072167.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 Query Match 16.5%; Score 506; DB 2; Length 568;  
 Best Local Similarity 29.3%; Pred. No. 2,8e-29;  
 Matches 168; Conservative 70; Mismatches 180; Indels 156; Gaps 27;  
 Oy 48 MFTVPLPIPPVQKRLVTNPVNGELIWEYVELEIKPTTHOV-----YPLDGSADLV 98  
 |||||  
 Db 29 MFTVPLPIPPVQKRLVTNPVNGELIWEYVELEIKPTTHOV-----YPLDGSADLV 98  
 |||||  
 Oy 99 GYVQMS-----PGPTFOVPGVEYVAFINN-----NAEAPNSVH 132  
 |||||  
 Db 77 AY-GTSSKATVPGPTLEIYVGVDTVYTMNHLPKSHLLPMDPTISPATPKHGIGPVAH 135  
 |||||

```

QY 133 LHGSESRFAADGWADITEPGSKD-----YYPNRSARLTMYHDHAMHITAENA 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 LHGHIHEPTSDGNA-DAMTAGFREREPKMTKTYLHTEKKQPGNMWYDHAMAGLTRVL 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 YGQAGLYLMDPA-EDALNPSGGEEDIPMLITSKOYTAGN-----NLVTNGELNSFW 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 LAGLVGAYILRRHAAVSPQLPTG-DEFDRPLIEFRSPFKGSITMNAATGNPSTIHPQM 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 -----GDVTHVNGQPMPEFKNVPERKREFRLDAAVSRSGLYFADDAIDTLRPFVISA 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 254 QPEYFGDVIIVNGKAMPRLNVRRRYRFRILNASNARFEKFEFSN-----GLDFTVGS 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 294 DGLLEHPADTSLIYISMAERYEVDFSDYAGKTIELNLGSGIGIGTDIDY-----347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 308 DSAVLSKPMVMTSILSPSEIYDVVDYFKSPSRIVLAN-----DAPYRPSGD 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 348 ---DNTDKVMRVVADDTQPTSVVPAANLRDVPSPPTTNPQRFGRGTGPTWINGV 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 PVNEGKVMKELI-NNESDDTCTIPKTL--INYPNADVSNAVLTRY-----ISMVEYV 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 AFADYQONRLAN-----VYV-----GTERMELINAGNWHPIHILVDFKVISRTS- 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 SNSDEPTHLVNGLYEADVYETPKSGTTEVWEVNLTRD-NHPLHILGLFRVVEQTL 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 453 -----GNN-----RTVMPYESGLKDVWL--GRRETVVEAH 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 LAGLEEFKECKMTKONDAVKQISKYARCKTAAYTAHEGKKNVFMGMGHYTRILIVRS 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 484 VA-----PF-----PGVYMFCHNLHEDHDM 506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 YIHFNASYPFDEPTOERG-YVYHCHILHEDHDM 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
G96734
spore coat protein-like protein, 24980-21957 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96734
R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Matzali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719
A:Accession: G96734
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-591 <STO>
A:Cross-references: GB:AF005173; NID:g6714313; PIDN:AAF26006.1; GSPDB:GN00141
C:Genetics:
A:Gene: F23N20.3
A:Map position: 1

```

Query Match 15.3%; Score 470; DB 2; Length 591;  
 Best Local Similarity 29.6%; Pred. No. 1.3e-26;  
 Matches 156; Conservative 66; Mismatches 183; Indels 122; Gaps 24;

```

QY 91 DUGSADLVGYDGM-----PGPTQVPRGVEVAFIN-----N 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84 DLPATPVFAV-STSRSATVAPGPTIAVGVDTYVWRNHLPLHMLPMDPTSPAIPKH 142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 AANPNSVHLHGSFSRAAFGMAED-----ITEPGS---FKDYIYPRKQARILMYDHAM 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 143 GGIPVYVHLHGHIHEPTSDGNADSWFTAGFKETGSKWTKTTHYVKKQPGNMWYDHAA 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 177 HTAENAYRGQAGLYMLT-DPAEDALNPSGGEEDIPMLITSKOYTAGN-----NLVTNN 231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 GLTRVALLAGLGLSGLYILRRSSVESPLRLPTG--REFRPLVIFDRSRKKGSIYMNATGN 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 GELNSFW-----GDVTHVNGQPMPEFKNVPERKREFRLDAAVSRSGLYFADDAIDTRL 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 PTHPQMOPEYFEDALIVNGKAMPRLTVRRKRYRFRITNASNARFEKFEFSN-----GL 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 287 PFVIVSDSGLEHPADTSLIYISMAERYEVDFSDYAGKTIELNLGSGIGIGTDD 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 DFTVGSDSAYLAKPVSTGSVLLAPSEIYDVLDVDFSKSTKATILNANAPYPSGDPTV 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 347 YDNTDKVMRVVADDTQPTSVVPAANLRDVPSPPTTNPQRFGRGTGPTWINGV 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 EENS-KVMKFIIT-NYKSEVDTSIPKTL--IEPRHNVSTSTFRITAEVYSSIDEPT 431
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 399 WT-INGVAFADYQONRLANP-VGIVE-----RRELINAGNWHPIHILVDFK 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 HLYINGLPY---NAVTEPKIGTSEFSKSLMLTIKAVWEVNLTRD-NHPLHILGLFR 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 447 V-----ISRTSGNNARTVMPYESGLKDVWL--GRRETV 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 VLEQTLVYKSEETIECKTRNDVAKCEISKYARCKTAAYTAHEGKKNVFMGMGHYTKI 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 479 VVEAHY-----APPGVYMFCHNLHEDHMAAFNATVL 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 LVRFSTIHNSNESYFPAIQOERG-YVYHCHILHEDHDMRPF-AMVL 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 5
AC0414
probable exported protein YP03409 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AC0414
R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0414
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-533 <KUP>
A:Cross-references: GB:AL590842; PIDN:CAC92639.1; PID:g1598136; GSPDB:GN00175
C:Genetics:
A:Gene: YP03409

```

Query Match 14.3%; Score 437.5; DB 2; Length 533;  
 Best Local Similarity 26.1%; Pred. No. 2.7e-24;  
 Matches 141; Conservative 70; Mismatches 209; Indels 121; Gaps 18;

```

QY 52 PLPIPEVKP-----RLTVNPNVNGGEIYWEYIKRFTHQVYDPLDSADLVGYDGSPGP 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 PLPIPLLPDPANGKTNL-NIQTGSYVWL-----PST-----ATQTWGNGALLDP 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 108 TPQVPRGVEVAFINNAEAPNSVHLHGSFSRAAFGMAEDITEPGSKFYDYYPNRSAR 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 AIRLQGRKAVTIDITNALPEATVTHHNGLEIPGEVGDGQFQALIQPAKROVFAVBOQA 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 168 TIMYHNDHAMHITAENAYRGQAGLYMLTDPAEADALNPSGGEEDIPMLITSKOYTAGN 227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 TCWFPHRTHSKTGHOVAMGLGLVLDSDSEFLPLPKQMGVDDIPVILQDK-----L 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 228 VTTNGELN-----SEWGDVIVHNGQPMPEFKNVPERKY-RFRFLDAAVSRSGLYE 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 LDKHGVDVQDVMYTAAGVFGDBMLTNGVYPV-QQITRGWRVLRLLGNCNARSLNLT-- 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 ADTDAIDTRLPEKVIASDSGLLEHPADTSLIYISMAERYEVDFSDYAGKTIELNLG 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 ----ALSQGRPMYVIASDSGLLAEPVYVRELPILMGERREVLDVTRD--GQSLDVLTVLPV 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY	337	SIGIGDPTDYNDDKVMRF---	VYADDTQDPDTSVYANLRDPV-----	378
		: : :   : :	: : :   : :	
Db	302	TOMGH-TLAPFDQPLVLRIPORSLAIGSQVLPELSVLVPELADTVGVDERKFOALMDPKL		360
		-----	-----	
QY	379	-----	-----	393
		-----	-----	
Db	361	DMLGMOALVARYGKAMAMNMNHDGMDGNRPDMGCKMGMHDGTNNGAPAFNFS		420
		-----	-----	
QY	394	RTGPTWTINGAAPADYONRLLANPVGTVVERMELINAGNGWTHPIHILVDFPKAISPTSG		453
		: : :   : :   : :   : :	: : :   : :   : :	
Db	421	HAN--RINGAKFSKTEPAF--DAKOGKYEKMTTISGEBDMILHPRHVHVGSTQFRIT----		471
		-----	-----	
QY	454	NNARTVMPYESGLKDYVL--GRRETVVVEAHH--APFPYVFEHCNLLIHEDDMMAAENA		511
		: : :   : :   : :   : :	: : :   : :   : :	
Db	472	ENGKPPAEHRRGMKDIYHVEGARSEILVFNFYLLAPASTPYAAHCHLLEHEDTGMMGLGFTV		531
		-----	-----	
QY	512	T 512		
		: : :   : :	: : :   : :	
Db	532	S 532		

## RESULT 6

Probable blue-copper protein yack precursor [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AC3582  
R:DelVecchio, V.G.; Kaparal, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,  
R.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzert, P.H.; Hagius, S.; O'Callaghan, D.; Levesque,  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3352; PMID:11756688  
A:Accession: AC3582  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-494 <RUP>  
A:Cross-references: GB:AE008918; PIDN:ALU53822.1; PID:g17984756; GSPDB:GN00191  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI10580  
A:Map position: 11

Query Match	13.1%	Score 403,	DB 2,	Length 494,
Best Local Similarity	25.4%	Pred. No. 8.3e-22,		
Matches 137,	Conservative 71,	Mismatches 214,	Indels 96,	Gaps 16;

```

0Y 52 PLRPIRQOPRLTIVNPNVNGOEIWEYIEKFTQVYVDDLSADLYVDXDSPPPTFOV 111
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10 PLRPLRPEVPPDS-----GIVRLKVKQGRIS-FAKGSTAASAGINCAYLGPLVRL 58
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 112 PRGVETVVRFINNA-EAPNSVHLGCSFSAAFEDGAEDITEPESKDYVYVPRQASRTLM 170
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 MSG-ESVYLSVENANADEFTLHMHGLFVPSHLDGPRHNVIAQWAMEPKVVAVQASFNW 117
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 171 YHDAAMHITAEANAYGQAGLYMLTDPADALNBSGCGEFDIPMLTISKQYANGNLV-- 228
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 118 FHPHLGHTARQAHNGIAGLMLVRDGKDAERGLPEYGYDDLPVLQDORR-VIEGDVAYA 176
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 229 -TTNSELMSFMCVDVHVGQOPWPFKNVPERKRFREFLDAAVSRSGLIAPADDAIDTLRPL 287
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 PDIMDLHGFRQDMLVINGALAPEARVPAAMVRLLLGANNRPHIRAP-----GRP 230
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 288 FKVIASDSGLLEHPADTSLYLSMAERYEVVDFSDYAGKTELENLGSGISIGITDXY 347
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 LTVIASDGSGLISQPSIEGLTISPERIEVLVDPSN--GEANDLVITYGDNGSGDDL----- 284
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 348 DNTDKVMREYVADDTTQPDTSVVPANLRDVPSPSTNTPRQFRF----- 392
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 ----HLMRPV--DPALEGRVAKPVPVSLDQPAAPDEKLVSQRRSPFFDERMAENKLMRQ 339
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 393 -----GRIGPWTING-----VAFADV---QNR 412
   |||::: : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      340 PSSNHAAGDDMDHDHMGSMAGMDHDMGSSAADACGALDALTSGVQMALADKPFDER    :| ||
OY      413 LIANPVGTVEEMELINANGWTHPIHLHYDEKYIIRSTGCNNATVPYESGLKDYYVL    :|| |
Db      400 IDVEAKLSWEITWEI--TSREMAHPFHIGASFRLTM---NGKRPAHOTGMKDALI    :|| |
OY      473 -GRRTVVVAHYAPFGVYMFCNLTHEDIDMAAF 509   :|| |
Db      454 DGAKEILVHFDRKAARSHPMFHCILLEHEVDGMAAQ 491 :|| |

```

## RESULT 7

Probable copper-binding protein yack - *Escherichia coli*  
 C:Species: *Escherichia coli*  
 C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
 C:Accession: G64735; S45200  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
 .A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of *Escherichia coli* K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: G64735  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Cross-references: 1-516 <BLAT>  
 A:Cross-references: GB:A0000121; GB:U00096; NID:g1786306; PIDN:AACT3234.1; PID:g17863  
 A:Experimental source: strain K-12, substrain M61655  
 R:Fujita, N.  
 submitted to the EMBL Data Library, January 1994  
 A:Reference number: S45181  
 A:Accession: S45200  
 A:Molecule type: DNA  
 A:Residues: 1-463, 'LIRAC', '470', 'IPLR', <FUU>  
 A:Cross-references: EMBL:D26562; NID:4473770; PIDN:BAAO5579.1; PID:Q1006123; PID:g4737  
 A:Experimental source: strain K-12, substrain M3110  
 C:Genetics:  
 A:Gene: yack  
 C:Keywords: copper

Query Match	13.1%	Score 401;	DB 2;	Length 516;
Best Local Similarity	28.8%	Pred. No. 1.2e-21;		
Matches 156; Conservative	58;	Mismatches 192;	Indels 86;	Gaps 17;

[illegible]

C:Genetics:  
A:Gene: yack

Experimental source: strain 0107:n/, substrate EBL933

A, experimental source: scldm v1.0/h/, subscldm EBL933

[illegible]

```

Query March 10.9% Score 33; DB 2; Length 474;
Best Local Similarity 28.1%; Pred. No. 1, 1e-16;
Matches 129; Conservative 50; Mismatches 212; Indels 68; Gaps 16;

QY 33 GLEKRVQAISBQYPMFTVPLRIPVKKPRLVTNPNVNGEIQEIYEVEIKPPTHQVYRDL 92
Db 19 GSLPLRAQASSSTQOPOQ-----LPVPLRLSR-----RGQPLFTL-----TLQRAHMAFGSN 63
QY 93 GSADLVYDGKMSPGCTFGVPRGVEVTVRFINNAEAPNSVHLHSGFSRAAFGMAEDIREP 152
Db 64 KKAALWVGINGVLTGTVTVAFVNGDDVKLLYSRLREPEVSMTISGLQVPETLWNGEARMIRP 123
QY 153 GSFKDYTPPNQASATLWYHDAHMHITTAENNAIRGQAGILMLTDPRAEDLNLNPSGIGEDI 212
Db 124 GEDWSPLYLPRVQOPRANCVYHANTPRRMAPHYHNGIAGCMWLVEDAVASKMPKPSHGVDVF 183
QY 213 PMILTSKYOTANGNLVTNGELNSFWGDVIVHNGQPMWPFKKVVEPRKYRFRFLDAVSRSF 272
Db 184 PLIIDKRLDNMGVEYENPRPAKGFVSGDTLLVNGAQSFVEVSRGWVRLRLILMNSNARY 243
QY 273 GLYFADTAITRLRPFKYIASDSGLIEHPADTSLITYISMAERYEVYPPFSDYACKTITELR 332
Db 244 TLQIUSD-----GRPLYVASDQGLPLAPVAVVQQLSLAPGRREVRIDMSQ--GGEVSI- 294

```





Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: G65088  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-470 <BLAT>  
A:Cross-references: GB:AE000384; GB:U00096; NID:92367186; PIDN:AACT6053.1; PID:91789394;  
A:Experimental source: strain K-12, substrain MG1655  
R:Coleman, J.  
Mol. Gen. Genet. 232, 295-303, 1992  
A:Title: Characterization of the *Escherichia coli* gene for 1-acyl-sn-glycerol-3-phosphat  
A:Reference number: S20459; MUID:92212294  
A:Accession: S20461  
A:Molecule type: DNA  
A:Residues: 1-89 <COL>  
A:Cross-references: GB:M63491; NID:9147296; PIDN:AAA24398.1; PID:9147299  
C:Genetics:  
A:Gene: *sufl*

Query Match 10.3%; Score 315.5; DB 2; Length 470;

Best Local Similarity 25.9%; Pred. No. 2,1e-15;

Matches 133; Conservative 71; Mismatches 235; Indels 75; Gaps 21;

QY 19 AVQASVPETSPATGHLERKVAQISFQYPMFTVPLPIPVKOPRLTVTPNVNGQETWYE 78  
DB 15 ALCAGAVPLKASAAQ-----OQPLVPPLLES-----RCQPL-FWT 51  
QY 79 VEIKPFTHQVYDLSADLVGDSGSPGTFQYRGVEYVRFINNAEAPNSVHLGSFS 138  
DB 52 VQ---RAHMSFPGTRASVWGIRGLPTIRWKGDDVKLISNRLTENVSMTVAGLOY 108  
QY 139 RAAPGMAEDITEPGSFKDYYPNROSARTLWYHDAMHITAENAYRGOAGLYMLTDPAE 198  
DB 109 PGLMGGRPARMMSPNADMAPVLPIRQNAATLWYHANTPNRTAQOYVNGLAGMMLVEDEVS 168  
QY 199 DALNLPSCYGEFDIPILTSKQYANGNLVTNGELNSFWGVYIHVNGOPWPKNVEPRK 258  
DB 169 KSLPIPNHGVDDFVYIIOCKRLDNFGTEPEYNEPGSGGVGDTLLVNGVQSPYEVSRCQ 228  
QY 259 YRFRFLDAVSRSGFLYFADTDALDTRLPFKYIASDSGLLEHPADTSLLYISMAERYEV 318  
DB 229 VRLRLINASNSRYOLOMND-----GRPLVHISGDQGLPAPVSVKOLSLAGERRETI 282  
QY 319 PDESDYAGKTIELRLNGSGISGIGTDTDYDNTDKVRFVAVD-----TQPDTSVY 370  
DB 283 VQMSN--GDEVSI--TCGEAASI-----VDRIREFPEPSSILVSTLVTLRP--TGEL 329  
QY 371 PANLRDVPFP-SPT---TTPRQFRFGRTGPTWTINGVAFADYQNRLLANVPVGTVERNE 426  
DB 330 PLVTDSLPRMLPTEIMASPIRSRDISLGDDPGINGQLM-DV-NRIDVTAQOGTWERM- 386  
QY 427 LINAGWTHPIHILVDKVISRTSGNNAATVMPY--ESGLKDVWL--GRRETVV--E 481  
DB 387 TVRADE--FOAHIEGVMTQI-----RVNNGAMPPEPDRGKMDIYVWDQGVELLVYFGQ 438  
QY 482 AHYAPPPGVYMFCHNLHEDHDMMAAFNATVLP 515  
DB 439 PSMAHFP--FYFNSQTFLEMAADRSGISIGOLLVNPYP 470

Search completed: October 3, 2002, 10:55:35  
Job time: 79 sec

[illegible]

PT Bilirubin oxidase prep. useful as an analytical enzyme - by  
PT culturing bilirubin oxidase in transformant culture  
XX  
XX  
PS Claim 1; Page 29-32; 32pp; Japanese.  
PS  
XX  
XX The sequence encodes bilirubin oxidase. The protein produced has a  
CC 38 amino acid signal peptide which is removed to give the mature  
CC protein (AAR0843).  
XX  
XX  
SQ Sequence 572 AA;

Query Match	100.0%;	Score 3067;	DB 14;	Length 572;
Best Local Similarity	100.0%;	Pred. No. 1.2e-270;		
Matches 572;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	MEKFTLGLAAALSLLENFNSNAOASVEPSTSATGHELFKRVQISPOXYMEFVPLRIDPYWQ	60
Db	1	mfxhtlgaaisllfnanaqaasvpretsapatghlfrvqeqispyqfmlfvrlprlrvpq	60
QY	61	PRLTVTNPNVNGEITWYXEVEIKRPTHQVPRDLSADLVGDKSPGPTFOVRPGVEVTNR	120
Db	61	prltvtvpnvngelwxyeveikrpthgvprdlsgadlvvgdkmspgptfvgvgetvnr	120
QY	121	FINNAEAPNSVHLNGSSRSRAFOGMAADIRERGSEFKQYXXVPRNQSARITLWHDHNAHITA	180
Db	121	finnaeapnsvhlngssrsrafdogmaadirergsefkdyypnrqsarclwyhdhahmhta	180
QY	181	ENAYRQAGLWMLDPAEDALNLPSCGGEFDIPMLITSKOYTANGNLVTNGELNSFWGD	240
Db	181	enayrqqaglymltdcpredalnlpsgygefdlpmlltskytangnlvtngelnsfwgd	240
QY	241	VIHWGQWPPEKNVPRKRYRFRFLDAASNSFGLEYADTDAIDTRLRPFKYIASDGLLEN	300
Db	241	vihwngqwpknavpepkyrrfrfdaasvnsfglyfadtaditrlrpfkyaasdglleh	300
QY	301	PADTSLYISAEERYEVVFPFSDYAGTIELRNIGSGISGIGDITDYDNTDKMKRRVAD	360
Db	301	padtslliyismaeeryevvfdtsdyagktelrnlngsigsigtdidntdkmkrrfvad	360
QY	361	DTTQPDNVSVPANLRDVPFSPPTNTPTROFRFGTGTWTINCVAFADVOYRLLANVPNG	420
Db	361	dtcpdsvvpaanlrdrvpsptncprqfrfgtgcvtlmgvafadvgnrllanvrpg	420
QY	421	TVERMELINAGNGTHTDIIHLVDKFYISRTSGNNATVPYRESGLKDVMWGLGRRETVVY	480
Db	421	tverwellingnagngtchpdlhlhldvfkvistrsgnartcmypsaglkdvwmjgretvvv	480
QY	481	EAAHTAPFPGVVMFCHNLIHEDHDMMAAFNATVLPDYGATATYFVDPMELLMQARPELG	540
Db	481	eahypfpgyvmfchnlhiedhmmaafnatvlpdygatatvfdvpmelmlqarpyelg	540
QY	541	EFOASOGGFSVOATERTIOTMAEERPPAADE 572	
Db	541	efqasgqfsqvavteritqmaeyrzppaade 572	

	RESULT	2
AAB81506		
ID	AAB81506	standard; Protein; 578 AA.
XX		
AC	AAB81506;	
XX		
DT	18-JUN-2001	(first entry)
XX		
DE	Billirubin oxidase.	
XX		
KW	Billirubin oxidase; phenol oxidising enzyme; phenol oxidase B;	
KW	detergent; paper production; pulp production; textile; food industry;	
	bleaching.	
XX		
OS	unidentified.	
XX		

PN	MO200121748-A1.	.
XX		.
PD	29-MAR-2001.	.
XX		.
PF	06-SEP-2000; 2000WO-EP08840.	.
XX		.
PR	22-SEP-1999; 99EP-0203120.	.
XX		.
PA	(UNIL ) UNILEVER NV.	.
PA	(UNIL ) UNILEVER PLC.	.
PA	(HIND-) HINDUSTAN LEVER LTD.	.
XX		.
PI	Convents D, Doornink M, De Vries CH, Wang H;	.
XX		.
DR	WPI: 2001-273462/28.	.
XX		.

Example 4, Fig 4; 46pp: English.

Example 4; Fig 4; 46pp; English.

The present sequence was used for comparison with the Stachybotrys chartarum phenol oxidising enzyme. The invention relates to detergent compositions comprising one or more surfactants and a phenol oxidising enzyme having at least 68% identity to the Stachybotrys chartarum phenol oxidising enzyme. Phenol oxidising enzymes may be used in the detergent paper, pulp, textile and food industries. They are used for preventing the transfer of dyes in solution from one textile to another during detergent washing, or in modifying the colour associated with dyes and coloured compounds having different chemical structures, such as in pulp and paper bleaching, bleaching the colour of stains on fabric and in detergent and textile applications.

50 Sequence 578 AA;

Query Match	65.5%;	Score 2010;	DB 22;	Length 578;
Best Local Similarity	65.4%;	Pred. No. 2.8e-174;		
Matches 383;	Conservative 71;	Mismatches 110;	Indels 22;	Gaps 10

[illegible]

OY	467	KOVWVLGGRETVVNEASHVAPPGVYWFHCHNLHHEHDMMAFNATVLDYGCNATVFDV	526
Db	473	Kdvvlarrevvyeahvafpfgvmlhcnhlhhehdmmaatnvclsdgyntlefid	532
OY	527	PMEEIWOARPYELGEFOASGOFVQAVTERTIOTMAEYRPAAD	572
		:   :   :   :   :   :   :   :   :   :   :   :	
Db	533	pmeplwrprrflgetengsgdlselatridrigemasfrpyagadd	578
	RESULT	3	
ID	AAB81505	standard; Protein; 583 AA.	
XX			
AC	AAB81505;		
XX			
DT	18-JUN-2001	(first entry)	
DE			
XX		Stachybotrys chartarum phenol oxidase B enzyme.	
XX			
XX		Stachybotrys phenol oxidase B: spob; phenol oxidizing enzyme;	
KW		detergent; paper production; pulp production; textile; food industry;	
KM		bleaching.	
XX			
OS		Stachybotrys chartarum.	
XX			
PN	WO200121748-A1.		
PD			
XX	29-MAR-2001.		
PF	06-SEP-2000; 2000WO-EP08840.		
XX			
PR	22-SEP-1999; 99EP-0203120.		
XX			
PA	(UNIL ) UNILEVER NV.		
PA	(UNIL ) UNILEVER PLC.		
PA	(HIND-) HINDUSTAN LEYER LTD.		
XX			
PI	Convents D, Doornink M, De Vries CH, Wang H;		
XX			
DR	WPI: 2001-273462/28.		
XX	N-PSDB: AAF82586.		
PT			
PT	New detergent compositions comprising a phenol oxidizing enzyme useful		
PT	in detergent or cleaning compositions, fiber treatment, processing,		
PT	finishing or production, paper and pulp production, or in starch		
PT	processing applications		
XX			
PS	Claim 1; Fig 2; 46pp; English.		
XX			
CC			
CC	The present sequence is a Stachybotrys chartarum phenol oxidizing		
CC	enzyme. The invention relates to detergent compositions comprising one or		
CC	more surfactants and a phenol oxidizing enzyme having at least 68%		
CC	identity to the Stachybotrys chartarum phenol oxidizing enzyme. Phenol		
CC	oxidizing enzymes may be used in the detergent, paper, pulp, textile and		
CC	food industries. They are used for preventing the transfer of dyes in		
CC	solution from one textile to another during detergent washing, or in		
CC	modifying the colour associated with dyes and coloured compounds having		
CC	different chemical structures, such as in pulp and paper bleaching,		
CC	bleaching the colour of stains on fabric and in detergent and textile		
CC	applications.		
XX			
XX			
XX			
Sequence	583 AA:		
XX			
XX			

QY	58	VKOPFLWTNPVNCQEIYYEVEELKPFHOVYPLDGSADLVGXGMSRGPTFOVPRCVER	117
Db	60	Ikapn - tvprnprrtgedlllyyemellstpnshqydpdpemmygyqgmpprpillvprgtes	118
QY	118	VYRFINNAE - APNSVHLHGSFSAEADGMAEDITTEPGSEKDYIYYPNROSARTLWYHDHA	175
Db	119	vrfvngsgentspsvnhlghsfstapfdgwaedcttgyeykdyuypprrgaarmlywjdha	178
QY	176	MHITENKAYRQAGLYMLTDPAPALNPSGSGEFDIPLMILTSKOYRANGNLVTINCELN	235
Db	179	msicteenymsgaqayymqdopaedalnpssygetdplvltcarrynaadgllstsgvcs	238
QY	236	SFMGDVIVHVGQMPWEPFKNVEPRKXFRFLDPAVARSGLYFADDAIDTTLPEKVIASDS	295
Db	239	sfmgvdlgvngqpwpmhnpvgrkryfrflnaavsrfsalyatseadsetrpfqvaiaadg	298
QY	296	GLEHPATSLILXYSMARREYVDFDSYAKCTELNLGSGIGITDIDYDNTOKVMR	355
Db	299	glllepvtdctclysmaerwevdfctsfagaidtlrnpba - dglvdepefntdctvmr	357
QY	356	FVVADDTIQPDTSVVPALRLRVPEP-----SPTNTNPQERFRGRTGPTWINGEFA	407
Db	358	fvdvdevlsepdteypanlrdvprfpeggnwqpnadd - etcfifgrangwtlmgvtfcs	415
QY	408	DYQKRLLANVPGYVERBELINAQNGRTHPHHLVLPKVISRISGNNARVPWDE - SGL	466
Db	416	dvenlltnvprdtveivrfleennsgwphvhlhldvrlvstst - -arvgeyeeaaql	472
QY	467	KDVWLGRRETVVVEAHARPPGVYMFCHLILIEDIDMMAAFNATVLPDYGNATYFVD	526
Db	473	kdvwlarrevvyeahvaprpyymhlchllinedidmmaafvctvlgdygnytefid	532
QY	527	PHEELWQARPELEBFOAQSGQSVQAVTEKTIQTMATERYPAADE	572
Db	533	pnehrwprpdllefengsgdfiselaatltdidigemafnayaqad	578

XX		RESULT 4
PI	Convents D, Doornink M, De Vries CH, Wang H;	AAB20097
XX		ID AAB20097 standard; Protein; 583 AA.
DR	WPI: 2001-273462/28.	XX
XX	N-PSDB; AAF82586.	AC AAB20097;
XX		XX
PT	New detergent compositions comprising a phenol oxidizing enzyme useful	DT 23-APR-2001 (first entry)
PT	in detergent or cleaning compositions, fiber treatment, processing,	XX
PT	finishing or production, paper and pulp production, or in starch	DE Stachybotrys chartarum phenol oxidase B.
PT	processing applications	XX
XX		KW Phenol oxidising enzyme; phenol oxidase B; spob gene; bleach;
PS	Claim 1; Fig 2; 46pp; English.	KW pulp; paper; textile; detergent.
XX		XX
CC	The present sequence is a Stachybotrys chartarum phenol oxidising	OS Stachybotrys chartarum.
CC	enzyme. The invention relates to detergent compositions comprising one or	XX
CC	more surfactants and a phenol oxidising enzyme having at least 68%	PV US6168936-B1.
CC	identity to the Stachybotrys chartarum phenol oxidising enzyme. Phenol	XX
CC	oxidising enzymes may be used in the detergent, paper, pulp, textile and	PD 02-JAN-2001.
CC	food industries. They are used for preventing the transfer of dyes in	XX
CC	solution from one textile to another during detergent washing, or in	PF 22-SEP-1999; 99US-0401476.
CC	modifying the colour associated with dyes and coloured compounds having	XX
CC	different chemical structures, such as in pulp and paper bleaching,	PR 22-SEP-1999; 99US-0401476.
CC	bleaching the colour of stains on fabric and in detergent and textile	XX
CC	applications.	PA (GENEV ) GENENCOR INT INC.
XX		XX
SQ	Sequence 583 AA;	PI Wang H;
		XX
		XX
Query Match	65.5%; Score 2010; DB 22; Length 583;	DR WPI: 2001-136715/14.
Best Local Similarity	65.4%; Pred. No. 2.9e-174;	DR N-PSDB; AAF30028, AAF20029.
Matches 383; Conservative 71; Mismatches 110; Indels 22; Gaps 10;		XX
OY 1 MEKHILGAALSL-LFNSSNAVOASVPV--ETSPATCHLEKFRVAOISPOVMFTVPLPDP'57		PT New phenol oxidizing enzyme, also useful in the detergent, paper and
I :   I  : :  I  :   : I I I I  :   I I I		PT pulp, textile or food industries, especially in modifying the colour
DG 1 misgaigavaglavigssvdarsvagrstlmpsg-ltkrtgtqslspalayeplpdp 59		PT associated with dyes and coloured compounds, as well as in anti-dye
		PT transfer applications
PS	Claim 1; Fig 2; 23pp; English.	



```

DB 421 gvsingvtfsdventenrlanvplgtvqlwqltnaagwctphihlvdfkilsrtgaatga 480
OY 455 NARVWAPYPS-ELKDVVWMLGRETVEVAHAPFPGVYMFHCHNLHEDHMAAANATV 513
DB 481 ttrgevpesaglkdvvytqkgevtveahyapwpyntfchmnlheddmmatnvtv 540
OY 514 LPDYGVNATVFDPMELMQARPYELGEFOAGSGFVSQVQATERTIQTMAEYRPAAD 571
DB 541 lpdgynstalcadpmdqdetrakpyvndvevrenatstdeieaqvlmsylpdyndp 598

```

RESULT 6

```

AA95538
ID AAY95538 standard; Protein: 627 AA.
XX
AC AAY95538;
XX
DE 10-OCT-2000 (first entry)
XX
DE Bipolaris spicifera phenol oxidizing enzyme.
XX
KM Phenol oxidizing enzyme; detergent; bleaching.
XX
OS Bipolaris spicifera.
XX
PN WO200039306-A2.
XX
PD 06-JUL-2000.
XX
PF 20-DEC-1999; 99WO-EP10287.
XX
PR 23-DEC-1999; 98US-0220871.
XX
PR 23-JUN-1999; 99US-0338723.
XX
PA (UNITL ) UNILEVER NV.
PA (UNITL ) UNILEVER PLC.
PA (HIND-) HINDUSTAN LEVER LTD.
XX
PI Bodie EA, Van Der Velden S, De Vries CH, Wang H;
XX
DR WPI; 2000-514528/46.
XX
DR N-PSDB; AAA50020.
XX
PT Detergent composition comprising novel phenol oxidizing enzyme obtained
PT from fungus or bacteria, useful for pulp and paper bleaching, bleaching
PT color of stains on fabric and for anti-dye redeposition
XX
PS Claim 8; Fig 3; 45pp; English.
XX
CC The present sequence is that of the Bipolaris spicifera
CC phenol oxidizing enzyme. The invention relates to detergent
CC compositions comprising novel phenol oxidizing enzymes that have at
CC least 60% identity with the phenol oxidizing enzyme of Stachybotrys
CC chartarum (see AAY95537), and which are obtained from a bacterium,
CC yeast or non-Stachybotrys fungus, especially B. spicifera.
CC Curvularia pallescens (see AAY95539) and Amersporium attum (see
CC AAY95540). The phenol oxidizing enzyme is capable of modifying the
CC colour associated with dyes or coloured compounds, and can be used
CC for pulp and paper bleaching, for bleaching the colour of stains on
CC fabric and for anti-dye transfer in detergent and textile
CC applications. It may also be capable of modifying the colour in
CC the absence or presence of an enhancer. Expression vectors and
CC host cells comprising a nucleic acid encoding a phenol oxidizing
CC enzyme, methods for producing the phenol oxidizing enzyme, and
CC methods for constructing expression hosts are provided.
XX
SQ Sequence 627 AA:

```

Query Match 58.7%; Score 1799; DB 21; Length 627;  
 Best Local Similarity 62.9%; Pred. No. 5,5e-155;  
 Matches 333; Conservative 67; Mismatches 123; Indels 6; Gaps 6;

```

OY 43 SPQPM-FTVPLPFPVKKOPRLVTNPNVGOETWYVEVEIKPETHHVPDLSADLVGD 101
DB 63 speyrlftrgplpripakepn-kltnvtnkelwyleivkprtgqvyslparlvayd 121
OY 102 GMSPGPTFQVPRGVETVVFRIINNAEAPNSVHLGFSRAAFDQWADIDIEPGSFKDYRP 161
DB 122 gslapgtllvprgteavvfingdreslhlhgsprsrpfdgwadmlmkgykkyrrp 181
OY 162 NROSARTLWTHDHAMHTTBNATRGQAGLXMLTDPEDALNLPSGGEFDIPMILTskoy 221
DB 182 nqgaarflyhdbamhvtlaenayfgagayiltcdpaedalglpsygykdydplvlskyy 241
OY 222 TANGNLVTNGELNSFMGVIHNGOPMPFKNVEPRKRYRFRFLDAVSRSGFYFADTQA 281
DB 242 naaglktsvgeksvsgdlhvnqgwpplnvepkyrlfllnaavsrnfalytkqdn 301
OY 282 IDTRLPFKVIASDSGLLEHPADTSLYISMAERYEVVFPDSYAGKTIELRNIGSGISGI 341
DB 302 tatrlpfgylaasagllthpvgtsdmvyaaeeyelvfdfapagqldlrlnf-akangi 360
OY 342 GTDPTDNTDKVWRVAVDDTQPTSVVPAHLRDVPPSPPTTNPBQRFGRGTFTWTI 401
DB 361 gtdddyantdkymrfhvsqtv-vdovsvpeqlsqifpadktdlhdhfrfrtngewri 419
OY 402 NGVAFADYONRLIANVPVGTVEERWELINANGWTHPIHILVDEKYSRSTSGNARTVMP 461
DB 420 nggldaventrlyakvprgtlwelelssgwsbphihlvfvravarydegtrvmp 479
OY 462 YE-SGLKDVVWMLGRETVEVAHAPFPGVYMFHCHNLHEDHMAAANATVLPDGYN 520
DB 480 yeaaglkdvvytqkgevtveahyapwpyntfchmnlheddmmatnvtvclqfygn 539
OY 521 ATY-FVDPMBELMQARPYELGEFOAGSGFVSQVQATERTIQTMAEYRPA 568
DB 540 etdfndepdprwsarptagdlrtarsglfseesirarvnelalegpy 588

```

RESULT 7

```

AA96762
ID AAY96762 standard; Protein: 627 AA.
XX
AC AAY96762;
XX
DE 09-OCT-2000 (first entry)
XX
DE Bipolaris spicifera phenol oxidizing enzyme.
XX
KM Phenol oxidizing enzyme; colour; dye; modification; detergent; stain;
KM pulp; paper bleaching.
XX
OS Bipolaris spicifera.
XX
PN WO200037654-A2.
XX
PD 29-JUN-2000.
XX
PF 20-DEC-1999; 99WO-US31009.
XX
PR 23-DEC-1999; 98US-0220871.
XX
PR 23-JUN-1999; 99US-0338723.
XX
PA (GENEW ) GENENCOR INT INC.
XX
PI Wang H, Bodie EA;
XX
DR WPI; 2000-452191/39.
XX
DR N-PSDB; AAA51315.
XX

```

New phenol oxidizing enzyme for modifying colors associated with dyes  
 or colored compounds, is obtained from fungus and is encoded by a  
 nucleic acid comprising a specific nucleotide sequence

PS Claim 8; Fig 3; 45pp; English.

CC This sequence is the *Bipolaris spicifera* phenol oxidizing enzyme.  
CC Phenol oxidizing enzymes encoded by nucleic acid sequences which  
CC hybridize to the *Stachybotrys chartarum* DNA (AA51313) are claimed, as  
CC long as the enzyme is capable of modifying the colour associated with  
CC dyes or coloured compounds. The enzymes are useful in detergent  
CC compositions and for modifying colors associated with dyes or coloured  
CC compounds which occur in strains in a sample. The enzymes are also useful  
CC for pulp and paper bleaching, anti-dye transfer in detergent and other  
CC textile applications.

CC Sequence 627 AA:

Query Match 58.7%; Score 1799; DB 21; Length 627;  
Best Local Similarity 62.9%; Pred. No. 5.5e-155;  
Matches 333; Conservative 67; Mismatches 123; Indels 6; Gaps 6;

QY 43 SPOYPM-FTVPLPPIPVKOPRLVTNPVNGOEIWEYVEIKFPTTHOVYDGLSADLVGYD 101  
DB 63 SPEYPLIFRQPIIPAKEPN-KLNPVNKELWYVEIKFPTGVSILPAILVGYD 121  
QY 102 GMSPGPTFOVPRGVEYVRFINNAEAPNSVHLGFSRAAFDGMADITEPGSFKDYYYP 161  
DB 122 GISPGLTIVPGTEAVRVFNGDRESSIHHSRSPRAPFDGWADMLMKQKDYYP 181  
QY 162 NQOSARTLWYHHAMHITENAYRGQAGLYMLTDPEDALNLPSSVGEEDIPMLTSKQY 221  
DB 182 NGAARFLWYHhnamvtaenayfgagaylltDpaedalglpsgykydpllvlskky 241  
QY 222 TANGNLVTNGELNSFWGVIVHNGOPMPFKVNEPKYRFRFLDAAVSRFSFGYFADTDA 281  
DB 242 NDDGLTGVSGEDKSWGDIHVNGPWFILNVEPKYLRFLNAAVSRNfalyfkqdn 301  
QY 282 IDTRLPEKVIASDGLLEHPRADTSLYISMAERYEVVFPDSYACTIELRLNGSIGI 341  
DB 302 TATRLPFGVIASDGLLTPVGTSDMYAAERYELVIFAPYAGQLTLRLNF-akangi 360  
QY 342 GTDTVDNTDKMRFVAVDDTQPTSVVPANLRDVPSPPTNPFROPRFGRTGPTWTI 401  
DB 361 GTDDYANTDKMRFVAVDDTQPTSVVPANLRDVPSPPTNPFROPRFGRTGPTWTI 419  
QY 402 NGVAFADVONRLLANVPVGTVERMELINAGCWTPHIIHLVDKVIISRTSGNNARTWAP 461  
DB 420 NGIGTADVONRLLANVPVGTVERMELINAGCWTPHIIHLVDKVIISRTSGNNARTWAP 479  
QY 462 YE-SGLKOVVHIGRETVVVEAHYARFPGVYMFHCHNLHEDHMAAFNATVLPDYGIN 520  
DB 480 YEaaglkdvwlgtrnetvveahyarpwdgvymfchhnlhedqdmmaadcvcklqifgyn 539  
QY 521 ATV-FVDPVEELMOARPYELGEFOAGSOFVQVATERIQTMAEYRPA 568  
DB 540 ECTDHDPEDPWRSARPTLAGDLTARSGLFSEESIRARVNEALALEGPS 588

RESULT 8

AA95539  
ID AA95539 standard; Protein: 627 AA.

AC AA95539;

DT 10-OCT-2000 (first entry)

XX Curvularia pallescens phenol oxidizing enzyme.

XX Phenol oxidizing enzyme; detergent; bleaching.

XX Curvularia pallescens.

OS WO200039306-A2.

XX 06-JUL-2000.

XX 20-DEC-1999; 99WO-EP10287.

XX 23-DEC-1998; 98US-0220871.

XX 23-JUN-1999; 99US-0338723.

XX (UNIT ) UNILEVER NV.

XX (UNIT ) UNILEVER PLC.

XX (HIND-) HINDUSTAN LEYER LTD.

XX Bodie EA, Van Der Velden S, De Vries CH, Wang H;

XX WPI; 2000-514528/46.

XX N-PSDB; AA50021.

XX Detergent composition comprising novel phenol oxidizing enzyme obtained

XX from fungus or bacteria, useful for pulp and paper bleaching, bleaching

XX color of stains on fabric and for anti-dye redeposition

XX Claim 8; Fig 10; 45pp; English.

XX The present sequence is that of the *Curvularia pallescens*

XX phenol oxidizing enzyme. The invention relates to detergent

XX compositions comprising novel phenol oxidizing enzymes that have at

XX least 60% identity with the phenol oxidizing enzyme of *Stachybotrys*

XX *chartarum* (see AA95537), and which are obtained from a bacterium,

XX yeast or non-*Stachybotrys* fungus, especially *C. pallescens*,

XX *Bipolaris spicifera* (see AA95538) and *Amerosporium atrum* (see

XX AA95540). The phenol oxidizing enzyme is capable of modifying the

XX colour associated with dyes or coloured compounds, and can be used

XX for pulp and paper bleaching, for bleaching the colour of stains on

XX fabric and for anti-dye transfer in detergent and textile

XX applications. It may also be capable of modifying the colour in

XX the absence or presence of an enhancer. Expression vectors and

XX host cells comprising a nucleic acid encoding a phenol oxidizing

XX enzyme, methods for producing the phenol oxidizing enzyme, and

XX methods for constructing expression hosts are provided.

Sequence 627 AA:

Query Match 58.5%; Score 1795; DB 21; Length 627;  
Best Local Similarity 63.1%; Pred. No. 1.3e-154;  
Matches 334; Conservative 63; Mismatches 126; Indels 6; Gaps 6;

QY 43 SPOYPM-FTVPLPPIPVKOPRLVTNPVNGOEIWEYVEIKFPTTHOVYDGLSADLVGYD 101  
DB 63 SPEYPLIFRQPIIPAKEPN-KLNPVNKELWYVEIKFPTGVSILPAILVGYD 121  
QY 102 GMSPGPTFOVPRGVEYVRFINNAEAPNSVHLGFSRAAFDGMADITEPGSFKDYYYP 161  
DB 122 GISPGLTIVPGTEAVRVFNGDRESSIHHSRSPRAPFDGWADMLMKQKDYYP 181  
QY 162 NQOSARTLWYHHAMHITENAYRGQAGLYMLTDPEDALNLPSSVGEEDIPMLTSKQY 221  
DB 182 NGAARFLWYHhnamvtaenayfgagaylltDpaedalglpsgykydpllvlskky 241  
QY 222 TANGNLVTNGELNSFWGVIVHNGOPMPFKVNEPKYRFRFLDAAVSRFSFGYFADTDA 281  
DB 242 NDDGLTGVSGEDKSWGDIHVNGPWFILNVEPKYLRFLNAAVSRNfalyfkqdn 301  
QY 282 IDTRLPEKVIASDGLLEHPRADTSLYISMAERYEVVFPDSYACTIELRLNGSIGI 341  
DB 302 TATRLPFGVIASDGLLTPVGTSDMYAAERYELVIFAPYAGQLTLRLNF-akangi 360  
QY 342 GTDTVDNTDKMRFVAVDDTQPTSVVPANLRDVPSPPTNPFROPRFGRTGPTWTI 401  
DB 361 GTDDYANTDKMRFVAVDDTQPTSVVPANLRDVPSPPTNPFROPRFGRTGPTWTI 419  
QY 402 NGVAFADVONRLLANVPVGTVERMELINAGCWTPHIIHLVDKVIISRTSGNNARTWAP 461  
DB 420 NGIGTADVONRLLANVPVGTVERMELINAGCWTPHIIHLVDKVIISRTSGNNARTWAP 479







```

XX 06-JUL-2000.
PD
XX 20-DEC-1999; 99WO-BP10287.
PF
XX 23-DEC-1998; 98US-0220871.
PR 23-JUN-1999; 99US-0338723.
XX
PA (UNITL) UNILEVER NV.
PA (UNITL) UNILEVER PLC.
PA (HIND-) HINDUSTAN LEVER LTD.
PI Bodie EA, Van Der Velden S, De Vries CH, Wang H;
XX
XX MPI: 2000-514528/46.
DR N-PSDB; AAS50018, AAS50019.
XX
PT Detergent composition comprising novel phenol oxidizing enzyme obtained
PT from fungus or bacteria, useful for pulp and paper bleaching, bleaching
PT color of stains on fabric and for anti-dye redeposition.
XX
XX Claim 2: Fig 5A-B; 45bp; English.
PS
XX The present sequence is that of a phenol oxidizing enzyme of
CC Stachybotrys chartarum MUC1.38898. A claimed detergent composition
CC comprises a phenol oxidizing enzyme having at least 60% identity to
CC the present sequence, and preferably obtained from a bacterium,
CC yeast or non-stachybotrys fungus, especially Bipolaris spicifera,
CC Curvularia pallescens or Alternaria alternata (see AAY9538-40). The
CC phenol oxidizing enzyme is capable of modifying the colour
CC associated with dyes or coloured compounds, and can be used for
CC pulp and paper bleaching, for bleaching the colour of stains on
CC fabric and for anti-dye transfer in detergent and textile
CC applications. It may also be capable of modifying the colour in
CC the absence or presence of an enhancer. Expression vectors and
CC host cells comprising a nucleic acid encoding a phenol oxidizing
CC enzyme, methods for producing the phenol oxidizing enzyme, and
CC methods for constructing expression hosts are provided.
XX
XX Sequence 594 AA:
SO
Query Match 56.2%; Score 1724.5; DB 21; Length 594;
Best Local Similarity 60.3%; Pred. No. 3.1e-148;
Matches 322; Conservative 75; Mismatches 128; Indels 9; Gaps 7;
QY 43 SPQYP-MFTVPLPIPVKQRLVTNPVNGQELWYEVETKPTTHOYVPLDGSADLVGD 101
DB 57 SPYNNLLYRNALPIPVKQPKMLITNPVNGQELWYEVETKPTTHOYVPLDGLVGD 116
QY 102 GMSPGPTFOVPRGCVETVREINNAEAPNSVHLHGSFSAFDDGMAWDITEPGSFKDYYP 161
DB 117 GMSPGPTFOVPRGCVETVREINNAEAPNSVHLHGSFSAFDDGMAWDITEPGSFKDYYP 176
QY 162 NROSATLWYHDAHMITAENAVRGAGLYMLTDPADALNLPDSGGEEDIPMILTSKOY 221
DB 177 NYGSARLLWYHDAHMITAENAVRGAGLYMLTDPADALNLPDSGGEEDIPMILTSKOY 236
QY 222 TANGNLTNGELNSFWGVYIHNGQWPKNRPKRRFRFEDAAVNSFGYFDTPA 281
DB 237 NADGTLTSTEGEDQDLWGDVILHNGQWPKNRPKRRFRFEDAAVNSFGYFDTPA 296
QY 282 IDTRLPFKYIASDSGLLEHPRADTSLYISMAERYEVVFDPSDAGRTIELNLGSGIGT 341
DB 297 PNVLRLFGYLAASDGLLQPVGVSNNLYLAERYEYLLIDFTNAGTDLIRNVAETND 355
QY 342 GIDTVDNMDKVRVAVADDTTQPTGSVPANLRDVPFESPTTNP--ROFRFGRTGPTW 399
DB 356 GDEDEYARLEVMRFVSSGTYE--DUSQVSTLRDVPF--PHKEGPAKDKIKFEISNGHY 413
QY 400 TINGVAFADYONLNLNVPYGTVERWELINAGGWHPIHIIHVDKRVISRTSGNNARV 459
DB 414 LINDVGLADVNERVLAKPELIVTEWELNSGGWSHPVHILVDIKLITKRGYR 471

```

```

QY 460 MPYES-GLKDVWLGRETVEVAHAYAPFPGVYMFCHNLHEDHDMMAFNATVLPDYG 518
DB 472 MPYESAGLKDVWLGRETVEVAHAYAPFPGVYMFCHNLHEDHDMMAFNATVLPDYG 531
QY 519 YNATFVDPMEELMWARPELGEFOASQGSVOAATERIQTMAEYRPPYAADE 572
DB 532 YLQEDFEDPMNPWRVAVPYRNDFNARGNFSAESLARTAYGELAEGEPYRIDE 585
RESULT 13
AAY96761
ID AAY96761 standard; Protein; 594 AA.
XX
XX AAY96761;
AC
XX 09-OCT-2000 (first entry)
DT
XX Stachybotrys chartarum phenol oxidizing enzyme.
DE
XX
XX Phenol oxidizing enzyme; colour; dye; modification; detergent; stain;
XX pulp; paper bleaching.
XX
XX Stachybotrys chartarum.
XX
XX WO200037654-A2.
XX
XX 29-JUN-2000.
XX
XX 20-DEC-1999; 99WO-US31009.
XX
XX 23-DEC-1998; 98US-0220871.
XX 23-JUN-1999; 99US-0338723.
XX
XX (GENV) GENENCOR INT INC.
XX
XX Wang H, Bodie EA;
XX
XX MPI: 2000-452191/39.
XX N-PSDB; AAS51313, AAS51314.
DR
XX
XX New phenol oxidizing enzyme for modifying colors associated with dyes
XX or colored compounds, is obtained from fungus and is encoded by a
XX nucleic acid comprising a specific nucleotide sequence
XX
XX Claim 1: Fig 5A-B; 45bp; English.
PS
XX This is the sequence of Stachybotrys chartarum phenol oxidizing enzyme.
XX Phenol oxidizing enzymes encoded by nucleic acid sequences which
XX hybridize to the coding DNA are claimed, as long as the enzyme is capable
XX of modifying the colour associated with dyes or coloured compounds. The
XX enzymes are useful in detergent compositions and for modifying colors
XX associated with dyes or colored compounds which occur in stains in a
XX sample. The enzymes are also useful for pulp and paper bleaching,
XX anti-dye transfer in detergent and other textile applications.
XX
XX Sequence 594 AA:
SO
Query Match 56.2%; Score 1724.5; DB 21; Length 594;
Best Local Similarity 60.3%; Pred. No. 3.1e-148;
Matches 322; Conservative 75; Mismatches 128; Indels 9; Gaps 7;
QY 43 SPQYP-MFTVPLPIPVKQRLVTNPVNGQELWYEVETKPTTHOYVPLDGSADLVGD 101
DB 57 SPYNNLLYRNALPIPVKQPKMLITNPVNGQELWYEVETKPTTHOYVPLDGLVGD 116
QY 102 GMSPGPTFOVPRGCVETVREINNAEAPNSVHLHGSFSAFDDGMAWDITEPGSFKDYYP 161
DB 117 GMSPGPTFOVPRGCVETVREINNAEAPNSVHLHGSFSAFDDGMAWDITEPGSFKDYYP 176
QY 162 NROSATLWYHDAHMITAENAVRGAGLYMLTDPADALNLPDSGGEEDIPMILTSKOY 221
DB 177 NYGSARLLWYHDAHMITAENAVRGAGLYMLTDPADALNLPDSGGEEDIPMILTSKOY 236

```



KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;  
KW fine chemical production; organic acid; proteinogenic amino acid;  
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
KW carboxylate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;  
KW diagnosis; Corynebacterium diptheriae; genetic engineering;  
KW Brevibacterium; environmental condition.  
XX  
OS Corynebacterium glutamicum.  
PN WO200100842-A2.  
PD 04-JAN-2001.  
XX  
PF 23-JUN-2000; 2000WO-IB00911.  
XX  
PR 25-JUN-1999; 99US-0141031.  
PR 08-JUL-1999; 99DE-1031636.  
PR 09-JUL-1999; 99DE-1032125.  
PR 09-JUL-1999; 99DE-1032126.  
PR 09-JUL-1999; 99DE-1032127.  
PR 09-JUL-1999; 99DE-1032128.  
PR 09-JUL-1999; 99DE-1032129.  
PR 09-JUL-1999; 99DE-1032226.  
PR 14-JUL-1999; 99DE-1032920.  
PR 14-JUL-1999; 99DE-1032922.  
PR 14-JUL-1999; 99DE-1032924.  
PR 14-JUL-1999; 99DE-1032930.  
PR 14-JUL-1999; 99DE-1032933.  
PR 14-JUL-1999; 99DE-1032935.  
PR 14-JUL-1999; 99DE-1032973.  
PR 14-JUL-1999; 99DE-1033002.  
PR 14-JUL-1999; 99DE-1033003.  
PR 14-JUL-1999; 99DE-1033005.  
PR 14-JUL-1999; 99DE-1033006.  
PR 31-AUG-1999; 99DE-1041378.  
PR 31-AUG-1999; 99DE-1041379.  
PR 31-AUG-1999; 99DE-1041390.  
PR 03-SEP-1999; 99DE-1041391.  
XX  
XX (BADI ) BASF AG.  
XX  
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
XX  
DR WPI, 2001-061974/07.  
XX  
XX N-PSDB; AAF71220.  
XX  
PT New isolated Corynebacterium glutamicum nucleic acid for production or  
PT modulation of production of fine chemicals such as amino acids,  
PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins,  
PT or enzymes -  
XX  
XX Claim 20; Page 373-374; 712pp; English.  
XX  
XX AAF7138 to AAF71357 encode the Corynebacterium glutamicum homeostasis  
CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The  
CC C. glutamicum HA genes (I) can be used in vectors for expression in host  
CC cells and production of fine chemicals, such as, an organic acid,  
CC proteinogenic or nonproteinogenic amino acid (preferred), purine or  
CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated  
CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,  
CC polypeptide or enzyme. The amino acids produced can be lysine, glutamine,  
CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,  
CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,  
CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can  
CC be modulated. The presence of (I) or HA proteins encoded by then are  
CC used for diagnosing the presence or activity of Corynebacterium  
CC diptheriae. (I) can be used to map the C. glutamicum genome or can be  
CC used as markers for genetically engineered Corynebacterium or  
CC Brevibacterium. The HA proteins encoded by the (I) are used to maintain  
CC homeostasis in C. glutamicum or help the microorganism to adapt to

CC different environmental conditions.  
XX  
SQ Sequence 497 AA:  
Query Match 16.1%; Score 494.5; DB 22; Length 497;  
Best local Similarity 30.9%; Pred. No. 3,2e-36;  
Matches 151; Conservative 67; Mismatches 219; Indels 51; Gaps 17;  
QY 53 LPIPPYKOPRLTYNTNVNGEITVEYEIKFTHOYPPIDGSDLYVGYGMSGPRFFQVP 112  
DB 30 Ippped-----lgtregssv-hialeqqtgesqllpdr-tktwgfnghlplrvlk 81  
QY 113 RGEIVYVRFITNAEAPNSVYLHGFSRAAFDGAEDITEPESKDYVYPRQSAFTLWYH 172  
DB 82 kgddvnydvinnidemttyvhgmklpaladgpbhpiipqgtwspwtvandaatlwyh 141  
QY 173 DRAMHTAENAYRGAGLYMLTDPADALNLPSCYGEFDIPMLITSKOYANGNLVTTNG 232  
DB 142 phthgltghayrglagmllvedeatkldlpreygvddipvlmchrfldegsldeedl 201  
QY 233 ELNSFWGDVYHVGOPWPFKNVPRKRYRFRFLDAVSRSGLYFADTDAIDTLPKRYA 292  
DB 202 pdlgyllgdrptlangitnahfdatttrvtrfvlnsgumrlynafsdtrt-----fgvla 255  
QY 293 SDSGLEHPADTSLYISMAREYEVDFSDYAGKTIELKNGSIG-GIGTD---TDYD 348  
DB 256 sdsglldepqdrtilalpggerweivele--pgevtlesvgyfednhyppddefypdlg 313  
QY 349 NTDKVRFFV--ADDTQPDTSVPANLBDVPSPPTTNTPRQFRFGRTGPTWTNGVA 405  
DB 314 msdsfgllltltpgsdaaq--apalpgvlykstepvidatertfflm-----tfsindlq 367  
QY 406 FAVYQRLANVPVGTVERKELINAGNGWHPHILVDFKVISRISGNNAKTVMEYEG 465  
DB 368 m-dmq-rvdvldhdqpevwivlndnsdwpfnfhvndarfkvI-klegld---vellfndg 421  
QY 466 LKDVWVL--GRETEVVE-AHYAPFGVYWFHCHNLIHEHDMAAFNMTVLDPGYCNAT 522  
DB 422 wkdtvllppatatlavefghypdpqpyymhcmlyhneogumngqf----- 468  
QY 523 VFVDPME 530  
DB 469 vlivepgde 476

Search completed: October 3, 2002, 10:56:41  
Job time: 145 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 11:02:16 ; Search time 19.26 Seconds

(without alignments)  
1149.926 Million cell updates/sec

Title: US-09-656-640A-4

Sequence: 1 MFKHTLGAALSLLENSNAV.....AVTERIQTMAYRYPYAADE 572

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3067	100.0	572	1	BLRO_MYRVE
2	713	23.2	513	1	COTR_BACSU
3	523	17.1	642	1	PHSA_STRAT
4	401	13.1	516	1	YACK_ECOLI
5	315.5	10.3	470	1	SUFI_ECOLI
6	311	10.1	470	1	SUFI_SALTY
7	228.5	7.5	520	1	LACI_TRAVI
8	228	7.4	311	1	SUFI_HAEIN
9	223.5	7.3	591	1	LACI_CRYPA
10	215	7.0	520	1	LACI_TRAVE
11	212	6.9	520	1	LACI_TRAVI
12	210	6.8	527	1	LACI_TRAVE
13	207	6.7	527	1	LACI_TRAVI
14	205.5	6.7	520	1	LACI_CORHI
15	205	6.7	519	1	LACI_TRAVE
16	205	6.7	609	1	COPA_PESM
17	205	6.7	624	1	PER3_CANAL
18	204	6.7	519	1	LACI_TRAVI
19	199.5	6.5	548	1	LACI_PHLRA
20	198.5	6.5	531	1	LACI_THACU
21	194	6.3	639	1	LACI_NEUCR
22	193.5	6.3	622	1	PER5_YEAST
23	191	6.2	605	1	PCOA_ECOLI
24	189.5	6.2	636	1	PER3_YEAST
25	189	6.2	533	1	LAC2_PLEOS
26	187.5	6.1	621	1	LAC2_PODAN
27	185.5	6.0	529	1	LAC1_PLEOS
28	178.5	5.8	619	1	LAC2_NEUCR
29	173.5	5.7	579	1	ASO_CUCMA
30	163.5	5.3	520	1	LACI_AGABI
31	163.5	5.3	520	1	LAC2_AGABI
32	162.5	5.3	622	1	YAKR_SCHPO
33	159.5	5.2	552	1	ASO_CUCPM

34	159.5	5.2	572	1	LAC3_THACU	002079	thanatephor
35	159	5.2	486	1	LAC1_BOMCI	012570	botrytis ci
36	150	4.9	587	1	ASO_CUCSA	P14133	cucumis sat
37	149	4.9	609	1	LAC1_EMBNI	P17489	emeritella
38	143	4.7	576	1	LAC1_THACU	P56193	thanatephor
39	141	4.6	473	1	LAC3_TRAVI	099049	trametes vi
40	134	4.4	599	1	LAC2_THACU	002075	thanatephor
41	131	4.3	462	1	CBSA_SULAC	054088	sulfolobus
42	113.5	3.7	578	1	ASO_TORAC	040588	nicotiana t
43	108.5	3.5	781	1	GCST_CADEL	019426	caenorhabdl
44	108	3.5	746	1	ABP_RAT	P36633	rattus norv
45	108	3.5	865	1	LOX2_ORYSA	P29250	oryza sativ

## ALIGNMENTS

RESULT	ID	BLRO_MYRVE	STANDARD	PRT	572 AA.
AC	012737				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Bilirubin oxidase precursor (EC 1.3.3.5).				
OS	Myrothecium verrucaria.				
OC	Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Myrothecium.				
OX	NCBI_TaxID=5532;				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	SPRAIN-MT-1;				
RX	MEDLINE=93366794; PubMed=8360171;				
RA	Koike S., Ando K., Kaji H., Inoue T., Murao S., Takeuchi K., Samejima T.;				
RT	Molecular cloning of the gene for bilirubin oxidase from Myrothecium verrucaria and its expression in yeast.;				
RL	J. Biol. Chem. 268:18801-18809(1993).				
CC	- FUNCTION: OXIDATION OF BILIRUBIN AND OTHER TETRAPYRROLES.				
CC	- CATALYTIC ACTIVITY: Bilirubin + O(2) = biliverdin + H(2)O.				
CC	- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR. CONTAINS 2 BLUE COPPER ATOMS PER MOLECULE.				
CC	- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.				
CC	- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL: D14081; BAA03166.1; -				
DR	EMBL: D12579; BAA02123.1; -				
DR	InterPro: IPR001117; Cu-oxidase.				
DR	Plant: PF00394; Cu-oxidase; 1.				
KW	Signal; Copper; Metal-binding; Oxidoreductase; Glycoprotein; Repeat.				
FT	SIGNAL	1	19		PROBABLE.
FT	PROPEP	20	38		
FT	CHAIN	39	572		BILIRUBIN OXIDASE.
FT	DOMAIN	98	194		PLASTOCYANIN-LIKE 1.
FT	DOMAIN	404	526		PLASTOCYANIN-LIKE 2.
FT	METAL	132	132		COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	134	134		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	172	172		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	174	174		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	436	436		COPPER (TYPE 1) (BY SIMILARITY).
FT	METAL	439	439		COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	441	441		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	494	494		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	495	495		COPPER (TYPE 1) (BY SIMILARITY).

```

RESULT 2
ID COTA_BACSU STANDARD; PRT; 513 AA.
AC COTA_BACSU P07788; C24818;
DT 01-AYG-1988 (Rel. 08, Created)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Spore coat protein A.
GN COTA OR PIG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168.
RX MEDLINE=97124186; PubMed=869499;
RA Borris R., Forwollik S., Schroeter R.;
RT "The 52 degrees-95 degrees segment of the Bacillus subtilis
chromosome: a region devoted to purine uptake and metabolism, and

```

RA containing the genes *cola*, *gabp* and *gua* and the *pur* gene cluster  
 RT within a 34960 bp nucleotide sequence";  
 RL Microbiology 142:3027-3031(1996).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RN STRAIN-168 / MARBURG;  
 RX MEDLINE=98116660; PubMed=9455482;  
 RA Kasahara Y., Nakai S., Ogasawara N., Yata K., Sadaie Y.;  
 RT "Sequence analysis of the *gROESL*-*cola* region of the *Bacillus subtilis*  
 RL genome, containing the restriction/modification system genes";  
 RN DNA Res. 4:335-339(1997).  
 RN [3]  
 RN SEQUENCE OF 1-37 FROM N.A.  
 RX MEDLINE=8801308; PubMed=2821284;  
 RA Donovan W., Zheng L., Sandman K., Losick R.;  
 RT "Genes encoding spore coat polypeptides from *Bacillus subtilis*.";  
 RL J. Mol. Biol. 196:1-10(1987).  
 RN [4]  
 RN SEQUENCE OF 1-34 FROM N.A.  
 RX MEDLINE=88286730; PubMed=3135411;  
 RA Sandman K., Kroos L., Cutting S.M., Youngman P., Losick R.;  
 RT "Identification of the promoter for a spore coat protein gene in  
 RL *Bacillus subtilis* and studies on the regulation of its induction at a  
 RT late stage of sporulation.";  
 RL J. Mol. Biol. 200:461-473(1988).  
 RN [5]  
 RN SEQUENCE OF 1-10 FROM N.A.  
 RP STRAIN-168;  
 RA Wray L.V., Person A.E., Fisher S.H.;  
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN BROWN PIGMENTATION DURING SPOREGENESIS.  
 CC -1- SIMILARITY: TO 5.ANTIBIOTICUS PHENOXAZINONE SYNTHASE (PISA).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: U51115; AAB62305.1; -.  
 DR EMBL: AB007638; BAA22774.1; ALT\_INIT.  
 DR EMBL: Z69107; CAB12449.1; -.  
 DR EMBL: X05678; CAA29165.1; ALT\_INIT.  
 DR EMBL: X07512; CAA30392.1; -.  
 DR EMBL: U31756; AAC44642.1; -.  
 DR PIR: A27393; A27393.  
 DR Subtilist; BG10490; *cola*.  
 DR InterPro; IPR001117; Cu-oxidase.  
 DR Pfam; PF00394; Cu-oxidase; 1.  
 KW Sporulation; Complete proteome.  
 FT CONFLICT 347 367  
 FT DESKRPT: LASYPSYOHRIQ -> TRAESRSTSPHRLYS  
 FT MKDT (IN REF. 1).  
 FT PTRGHH -> RHAHEHL (IN REF. 1).  
 FT GRAVPRP -> VRCPRAA (IN REF. 1).  
 SO SEQUENCE 513 AA; 58499 MW; 836B83BA58D75F87 CAC64;

```

Query Match          23.2%; Score 713; DB 1; Length 513;
Best Local Similarity 33.1%; Pred. No. 3.5e-44;
Matches 177; Conservative 77; Mismatches 179; Indels 102; Gaps 16;

QY 49 FTVPRIP-----PVQPRRLVTNPVNGOEIVWEIKRPTHQVYPDLSADLVGYDGS 104
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6 FVDALRIPRLPKVQO-----SKETIYEVTMECTHQLRDLPTPLMGYINGLF 55

QY 105 PGRTEQVPRGVEYTVRFINN-----AEPNSVYHLGSESRAAF 142
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 56 PGRTEVKKRNNVYVKKMNLPSHPLRLPDTHIHSDSOHEPEVKYTVHLLGGVTPRDS 115

QY 143 DGMAE-----DITEPGSF--KDYVYPNROSARTLWYHDMHAMHTAENAYRGOAGLYMLT 194
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```



DB 116 DGYPEAMFESQDEPQGTGTYFREYVYHHPNQRGAILMYHDHMAALITLNYAGLVAYIIH 175  
QY 195 DPAEDALNLPSCYGEEDIPMLITSKOYTANGNLVTNGEIN-----SFMGDVI 242  
DB 176 DPEKRLKLPDS--DEVDYPLILITDRITNEDGSLFYPSAPSPSPSLPNPSTVAFCCETI 233  
QY 243 HVGQFMPFKNVPRKRYREFELDAVSRFGYFADTDAIDTRLPEKVIASDGLLEHFA 302  
DB 234 LVNGKWPLYLEVPKRYKFRVIVNASMTRYNL-----SLDNGGDFLQISDGLLPRSV 287  
QY 303 DTSLITISMAEREVYVDFSDYAGKITELRNLGSGISGIGITDIDYNTDKVAFVADDT 362  
DB 288 KLSFSLAPAEKRDIIIDFPAYEGESILIASAGCGDVNPEDAN-----IMQFRYTKPL 343  
QY 363 TOPDTSVPAFLRDVPPSPPTTNTPROFR-----FGRTGPTWTINGVAFADVQNR 412  
DB 344 AOKDESKRKYLL--ASYPSQHERIQNIRKLKLAGTODEGR--PULLLNKRWHDP--- 336  
QY 413 LIANVP-VGTEVERMELINAGNGWTHPIHLVDPKVISRTSGNNAR----- 457  
DB 397 -VTEPKVGTEIWSIINPTRG--THPIHLVSPRVLDRRPDIARQESGELSYTGPAV 454  
QY 458 TWNPYFSGKLDVWVGRRREVVEAHYAPPYGMCHMLIHEDHMAAFNAT 512  
DB 455 PPPSEKGMKDTIQAHAGEYLRIATFPGYSGRYVWHCHLHEHDYDMRPMNDIT 509  
RESULT 3  
PHSA\_STRAT STANDARD: PRT: 642 AA.  
ID PHSA\_STRAT 053692;  
AC 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Phenoxazinone synthase (EC 1...5) (PHS).  
GN PHSA.  
OS Streptomyces antibioticus.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OY NCBI\_TaxID=1890;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.  
RX STRAIN=IMRU 3720;  
RC MEDLINE=96011355; PubMed=7592317;  
RA Hsieh C.-J., Jones G. H.;  
RT "Nucleotide sequence, transcriptional analysis, and glucose  
regulation of the phenoxazinone synthase gene (phsa) from  
Streptomyces antibioticus."  
RT J. Bacteriol. 177:5740-5747(1995).  
RN [2]  
RP SUBUNIT.  
RX MEDLINE=82066838; PubMed=7305384;  
RA Choy H.A., Jones G.H.;  
RT "Phenoxazinone synthase from Streptomyces antibioticus: purification  
of the large and small enzyme forms."  
RT Arch. Biochem. Biophys. 211:55-65(1981).  
CC -1- FUNCTION: CATALYZES THE LAST BUT TWO STEPS IN THE PUTATIVE  
BIOSYNTHETIC PATHWAY OF ACTINOMYCIN.  
CC -1- CATALYTIC ACTIVITY: 4 4-methyl-3-hydroxyanthraniloyl pentapeptide  
+ 3 O(2) = 2 actinomycin acid + 6 H(2)O.  
CC -1- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH  
CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE  
2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).  
CC -1- SUBUNIT: HOMODIMER (SMALL FORM) OR HOMOHETEROMER (LARGE FORM).  
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.  
CC -1- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.  
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation-  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL: U04283; AA06668.1; ALT\_INIT.  
DR InterPro: IPR001117; Cu-oxidase.  
DR InterPro: IPR002355; Multicopper oxidase2.  
DR Pfam: PF00394; Cu-oxidase; 1.  
DR PROSITE: PS00079; MULTICOPPER OXIDASE1; 1.  
DR PROSITE: PS00080; MULTICOPPER OXIDASE2; 1.  
KW Oxidoreductase; Repeat; Metal-binding; Copper;  
KW Antibiotic biosynthesis.  
FT INIT\_MET 0  
FT DOMAIN 86 223  
FT METAL 161 161 PLASTOCYANIN-LIKE 1.  
FT METAL 163 161 PLASTOCYANIN-LIKE 2.  
FT METAL 163 161 COPPER (TYPE 2) (BY SIMILARITY).  
FT METAL 201 201 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 203 201 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 524 524 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 527 527 COPPER (TYPE 2) (BY SIMILARITY).  
FT METAL 529 529 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 602 602 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 603 603 COPPER (TYPE 1) (BY SIMILARITY).  
FT METAL 604 604 COPPER (TYPE 3) (BY SIMILARITY).  
FT METAL 608 608 COPPER (TYPE 1) (BY SIMILARITY).  
FT METAL 613 613 COPPER (TYPE 1) (BY SIMILARITY).  
SQ SEQUENCE 642 AA; 70113 MW; E0B39C0BA3364E48 CRC64;

Query Match 17.18; Score 523; DB 1; Length 642;  
Best Local Similarity 27.8%; Pred. No. 2.5e-30;  
Matches 176; Conservative 83; Mismatches 199; Indels 174; Gaps 28;

QY 27 ETSPATGHLFKRAQISPOYPMFTVPIPIPVYQPRILTVNPNVNGSITWYEYKFTH 86  
DB 33 EQAPAPP-----ETPP-----FAAPLTVPVLP--ASDEVTRT-----ETALRPTWV 74  
QY 87 QVVPDLGSADLVGYDGKSPGPTFOVPGVEYVRFIN-----NAE 126  
DB 75 RLHPQLPPTLMKGYDGVPEPTIEVRGQVRIAMTRIKGSEYPTVSVEPLGPGTTP 134  
QY 127 APNS-----VHLGSFSAFAGMAEDITTEPSFKDYYPNROS 165  
DB 135 APMTPEGRGVEPNKDVAALPAMSVTHLGAQOTGGNDGMADNAVGGDAQLSYPRDHO 194  
QY 166 ARTLWYHDHAMHTTAENAYGQAGLWLTDPADALNLPSCYGEEDIPML-----TSK 219  
DB 195 ATQWVYHDHAMITRTMVMAGLYGTLYVRDDEBDALGLPS--GDREIPLIADARNLDTDE 252  
QY 220 QYRANGNLY-----TTNGELMS-----FMGVYIHVNGSPMPFKNVPRKRYREFELDA 267  
DB 253 DGRINGRLKRLKTYIVQGSNPEFGKPVSIPEFGPYTYVNGRIMPVADDDGMYRLRVNAS 312  
QY 268 VSRSGLYFADTDAIDTRLPEKY--IASDSGLEHNP-----DT-SLITISMAEREVVF 319  
DB 313 NARIYNVLIDED--DRPVGVVHQIGSDGLLPRVPVDFDTLPLVLSAAPAREFDLIV 370  
QY 320 DESDYAGKITELNGL-----GSIGIGITPTDIDNTDKVAFVADDTTPDTSVVP 371  
DB 371 DFLAAGRRRLRLVDKGGAPGAPTPDPLGV-----RYPEVMEFRVR--ETCEDSDPALP 422  
QY 372 ANL-----RDVPF-----PSPTNT-----PROFRF----- 392  
DB 423 EVLSGSRFRRMSHDIPIGHRLIVLTPTGKSGSGHPEIWEAEVEDPADVOVPAEGVIQVT 482  
QY 393 GRTGPTWTINGVAFADVQNRILANVPVGTVERMELINAGNGWTHPIHLVDPKVISR-- 450  
DB 483 GAGGRRTYRRTA--ATPNOLGFTIGCTHEQWTFNL--SPIAHMHIHLADQVIGRDA 540  
QY 451 --TSG-----NMART-----VMYFESGLKDVWVGRRREVVEAHYAPPYGM 492  
DB 541 YDASGFDALGTRTVPRLDPDTPVPLAPNELCHKDVFQVPGQGLRVKMGCFDAGVGRFM 600  
QY 493 FCHGNLIHEDHMAAFNATVLP-----DYG 518

DB 601 YHCHLEHEDMGMRPF--VVMPEALKEPDHG 630

RESULT 4

YACK\_ECOLI STANDARD: PRT: 516 AA.

AC P36649; P75655;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable blue-copper protein yack precursor.

GN YACK OR B0123.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=562;

RN NCBI\_TaxID=562;

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / W3110;

RA MEDLINE-94261430; PubMed-8202364;

RA Fujita N., Mori H., Yura T., Ishihama A.;

RT the 2.4.4.1 min (110,917-193,643 bp) region.";

RL Nucleic Acids Res. 22:1637-1639(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RA MEDLINE-97426617; PubMed-9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP SEQUENCE OF 29-40.

RC STRAIN-K12 / EMG2;

RA MEDLINE-97443375; PubMed-9298646;

RA Link A.J., Robison K., Church G.M.;

RT "Comparing the predicted and observed properties of proteins encoded

RL in the genome of Escherichia coli K-12.";

RN Electrophoresis 18:1259-1313(1997).

RN [4]

RP IDENTIFICATION BY MASS SPECTROMETRY.

RA MEDLINE-99420866; PubMed-10493123;

RA Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;

RT "Enrichment of low abundance proteins of Escherichia coli by

RL hydroxypapillate chromatography.";

RC Electrophoresis 20:2181-2195(1999).

RL -I- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH

CC CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE

CC 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).

CC -I- SUBCELLULAR LOCATION: Periplasmic (Potential).

CC -I- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.

CC -I- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.

CC -I- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A

CC FRAMESHIFT IN POSITION 464.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: D26562; CAB20297.1; ALT\_FRAME.

CC EMBL: AB000121; AAC73234.1; --

CC PIR: S45200; S45200.

CC Ecogene: BGI2318; yack.

CC InterPro: IPR001117; Cu-oxidase.

CC InterPro: IPR002355; MultiCu\_oxidase2.

DR Pfam: PF00394; Cu-oxidase; 1.

DR PROSITE: PS00080; MULTICOPPER\_OXIDASE2; 1.

KW Periplasmic; Signal; Copper; Metal-binding; Repeat; Oxidoreductase;

KW Complete proteome.

FT SIGNAL 1 28

FT CHAIN 29 516

FT DOMAIN 67 163

FT DOMAIN 164 410

FT DOMAIN 411 516

FT DOMAIN 101 101

FT METAL 103 103

FT METAL 141 141

FT METAL 143 143

FT METAL 443 443

FT METAL 446 446

FT METAL 448 448

FT METAL 499 499

FT METAL 500 500

FT METAL 501 501

FT METAL 505 505

FT METAL 510 510

FT METAL 510 510

SO SEQUENCE 516 AA; 56556 MW; 37D96B1C31CF30B CRC64;

Query Match 13.1%; Score 401; DB 1; Length 516;

Best Local Similarity 28.8%; Pred. No. 1.2e-21;

Matches 136; Conservative 58; Mismatches 192; Indels 86; Gaps 17;

99 GYNGMSPGPRFYQVGRVEYVVRFINNAEAPNSVHLGSSRAAFDDMAEDITPPGSFKDY 158

DB 68 GYGNLLGPAVKLKQRAKAVTDVLDNLTETTLHMHGLEVEGSDGPGGIIIPPGKRSV 127

QY 159 YYPNROSARTLWYHDAMHTAENAYRGOAGLYMLDPAEDALNLPSCGGEFDIPILTS 218

DB 128 TLNVDDPRAATCFPHRPHQHGKTGRQVAMGLAVYIEDDELKMLRKQMGIDVPIYOD 187

QY 219 KOYTANGNLVTNGELN-----SFWDVIVHNGQWPFRKNVPRKY-RFRFLDAAVSR 271

DB 188 KRSADGQI---DYQLDVMTAAVGMFEGDTLLTGALYIP-QHAPRGMRLRLILNCNARS 243

QY 272 FGLYFADTDALDRLPFKYIASDSGLLEHRAVDLSLYISAEKVEVFPESVDAGTIEL 331

DB 244 --LNFATSD---NRPLVYIASDGLLPPEKVSLEPLHGEFEVLEVND--NRPFD 295

QY 332 RNLGSGIGIGTDTDDNDIKVMR---FVYADDTQPDV-SVPAVLRDVPFSPPTNT 387

DB 296 VILPVSGMGKAI-APDKRPHVRIPIAISAGALPDLTSLPA-----LPSLEGLTV 348

QY 388 RQFR-----FGRTG-----PTW 399

DB 349 RKLQSLMDPMLDMGMQIMLEKYGDQAMAGMDHSGMHGHNMMHNGKFEFNHAN 408

QY 400 TINGVAFADYQNLNLANVGVYERWELINAGMGWPHRIHLVDRKVISRTSGNARTV 459

DB 409 KINGQAFD--MKNPMPAAKGGYERNAVIGVGDMILHPRHIGTORFILS---ENKRP 462

QY 460 MPYESILKDVVWL--GRRETVVE-AHYAPFPGVYMFHCHNLIEDHDDMAAF 509

DB 463 AAHRAGKDTVKEGVNSEVLYKFENDAKREHAYMAHCHLLEHEDMGMLGF 514

RESULT 5

SUPL\_ECOLI STANDARD: PRT: 470 AA.

AC P26648;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Protein sufl precursor.

GN SUFL OR B3017.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Berna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE OF 1-89 FROM N.A.  
 RX MEDLINE=92212294; PubMed=1557036;  
 RA Coleman J.;  
 RT "Characterization of the *Escherichia coli* gene for  
 RT 1-acyl-sn-glycerol-3-phosphate acyltransferase (plac).";  
 RL Mol. Genet. 232:295-303(1992).  
 RN [3]  
 RP SEQUENCE OF 28-39.  
 RC STRAIN=K12 / EMG2;  
 RX MEDLINE=97443975; PubMed=9298646;  
 RA Link A.J., Robison K., Church G.M.;  
 RT "Comparing the predicted and observed properties of proteins encoded  
 RT in the genome of *Escherichia coli* K-12.";  
 RL Electrophoresis 18:1259-1313(1997).  
 CC -1- FUNCTION: INVOLVED IN CELL DIVISION, SUPPRESSES A FTSI MUTATION.  
 CC -1- SUBCELLULAR LOCATION: Periplasmic.  
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U28377; AAA69185.1; -  
 DR EMBL: AE000384; AAC76053.1; -  
 DR EMBL: M63491; AAA24398.1; -  
 DR PIR: S20461; S20461.  
 DR Ecogene: EG11376; sufi.  
 DR InterPro: IPR001117; Cu-oxidase.  
 DR Pfam: PF00394; Cu-oxidase; 1.  
 DR Periplasmic; Signal: Complete proteome.  
 FT SIGNAL 1 27  
 FT CHAIN 28 470 PROTEIN SUFI.  
 FT DOMAIN 68 164 PLASTOCYANIN-LIKE.  
 FT SEQUENCE 470 AA; 51858 MW; C843A5A4CB146688 CRC64;  
 SQ  
 Query Match 10.3%; Score 315.5; DB 1; Length 470;  
 Best Local Similarity 25.9%; Pred. No. 1.6e-15;  
 Matches 133; Conservative 71; Mismatches 235; Indels 75; Gaps 21;  
 OY 19 AVOASVPEPTSPATGHLFRVAQISQVPMFYVLPPIPKVQRLVTVNPGCELTWYE 78  
 DB 15 ALCAGAVPLKASAGQ-----QQLPVPPLLEER-----RGQPL-FMT 51  
 OY 79 VEIKPTHOVYDPLGSAADLVGYDGMSPGTFVPGVETVREINNAEPNSVHLGSFS 138  
 DB 52 VQ---RAHMSFPGRTRASVWINGRYLPTIRVWKGDDVKLIYSNRLTEVSWTVAGLQY 108  
 OY 139 RAAFDGMADIDEPGSKFYIYPNROSARTLWYHDHAMITAEANVRGQAGLYMLTDPAR 198  
 DB 109 PGRLMGAPRPMSPNDAMPVLPPIRQNAATLWYHANTPRTAQOYVNGLAGMVLVEDEVS 168  
 OY 199 DALNPSGGEEDIPMILTSKOTYTAGNLVTTNGELNSFMGDVHYHNGDPWPKVNEPRK 256  
 DB 169 KSLPIFNHGVDDFYIYIDKRLDNGCTPEYNEPSSGSGVGGDLYLVNGQSPVVEYSRGM 228  
 OY 259 YRFRFLDAAVSRSGFLYFADTDAIDTRLPFKVYIASDSGLLEHPADTSLLYISMAERYEVV 318

DB 229 VRLRLNANSSRRYQLOMND-----GRPLHVISGDOGFLPAPSVKQSLACGERREIL 282  
 OY 319 FDFSVDYAGKTELRLNGLSGTIGTIDTDYNTDKVRFVYADD-----TQOPTSVY 370  
 DB 283 VDMNS--GDEVSI--TCGEAAST-----VDRIKFFEPSSILVSTLVTLRP-TGL 329  
 OY 371 PAMLRVPPF-SPT---TNPPOFRGRCPTWTINGVAPADVQNRLLANVPYGVVERME 426  
 DB 330 PLVTDLSLPMKPLFTELMAQSPITRSISLGDDEINGQLM-DV-NRIDVTAQOGTWERW- 386  
 OY 427 LINAGMGWTHPIHILVDFKVISRTSGNNAITVMPY--ESGLDYYWL-GRETVVYV--E 481  
 DB 387 TVRADE--PQAFHIEGVMEQI-----RNVNGAMPFEDRGMDYVWDGQVELLYVFGQ 438  
 OY 482 AHYAPFGYVFMFCHNLHEDHDMAAFNATVLP 515  
 DB 439 PSMHAPP--FYFNSQTLMAADRGISQGLLVNPPV 470  
 RESULT 6  
 SUFI\_SALTY STANDARD: PRT; 470 AA.  
 ID SUFI\_SALTY  
 AC P40799;  
 DT 01-FEB-1995 (rel. 31, Created)  
 DT 01-MAR-2002 (rel. 41, last sequence update)  
 DT 01-MAR-2002 (rel. 41, last annotation update)  
 DE Protein sufi precursor.  
 GN SUFI OR STM3172 OR STY3349.  
 OS *Salmonella typhimurium*, and  
 OS *Salmonella typhimurium*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Salmonella*.  
 OX NCBI\_TaxID=602, 601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S. typhimurium; STRAIN=LT2;  
 RA Cong J., Schmid M.B.;  
 RL Submitted (APR-1994) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S. typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of *Salmonella enterica* serovar *Typhimurium*  
 RT LT2.";  
 RL Nature 413:852-856(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S. typhimurium; STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagsels K.,  
 RA Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*  
 RT enterica serovar *typhimurium* CT18.";  
 RL Nature 413:848-852(2001).  
 CC -1- FUNCTION: INVOLVED IN CELL DIVISION, SUPPRESSES A FTSI MUTATION  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Periplasmic (by similarity).  
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 413.





NW	Glycoprotein; Repeat.			POTENTIAL.
FT	SIGNAL	1	20	LACCASE.
FT	CHAIN	21	591	PLASTOCYANIN-LIKE 1.
FT	DOMAIN	66	189	PLASTOCYANIN-LIKE 2.
FT	DOMAIN	198	356	PLASTOCYANIN-LIKE 3.
FT	METAL	416	551	COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	126	126	COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	128	128	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	171	171	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	173	173	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	463	463	COPPER (TYPE 1) (BY SIMILARITY).
FT	METAL	466	466	COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	468	468	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	533	533	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	534	534	COPPER (TYPE 1) (BY SIMILARITY).
FT	METAL	535	535	COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	539	539	COPPER (TYPE 1) (BY SIMILARITY).
FT	METAL	?	?	COPPER (TYPE 1) (BY SIMILARITY).
FT	CARBONYD	121	121	N-LINKED (GLONAC. .) (POTENTIAL).
FT	CARBONYD	234	234	N-LINKED (GLONAC. .) (POTENTIAL).
FT	CARBONYD	242	242	N-LINKED (GLONAC. .) (POTENTIAL).
FT	CARBONYD	265	265	N-LINKED (GLONAC. .) (POTENTIAL).
FT	CARBONYD	323	323	N-LINKED (GLONAC. .) (POTENTIAL).
FT	CARBONYD	407	407	N-LINKED (GLONAC. .) (POTENTIAL).
FT	CARBONYD	425	425	N-LINKED (GLONAC. .) (POTENTIAL).
QO	SEQUENCE	591 AA:	64696 MW;	B2F44CB3AADA07701 CRG64:

QY		9	AALILEFNSNAVOASP-----VPETSPATGHLERVAQLSPOYPMTF---VPLPIP-PV	58
Dd		7	ALFSGLLASOLSNAPALSIHPLEBRQPCNCNTAENRACMWISGSIDITTDVEVKRPRLGV	66
QY		59	KOPRLVTNPVNGCEIWEYEVEIKPFTHQVPPD-LGSADLVGYDGMSPGPTFOVPGRVFH	117
Dd		67	ROYDLTLTLOAEN-----W-----LCPDPGEVEDMALVGNGLGPVIAHOMODTI..	110
QY		118	VVRFINNAAE-APNSVHLHG--SFSSRAAFDGAEDITE-----PSFSKDYXYPNNOSARTL	165
Dd		111	SVYTATNLTKYTGCTITHMHGIROLTFNIODE-VNGSITECPLIPPNNGSKSTYFIHAHQY-TS	168
QY		170	MYHDHAMHTTAENAYRGCAGLMYLTDPDAEDALNLPSCGYEFDI---PMILTSKOYTANGN	226
Dd		169	MYSH--HFSAOYG-NGIVGAIQIDGPA---SLP-----YDIDLGPLVLSDYYKYAKDE	215
QY		227	LVTINGELNLSRWGDVIHVNGQPWFPKVNEP-----KKYPRFLDAAVSR	270
Dd		216	LVVYTOQNAPPASDNVLFENG----TNINPANTTOGOYKTITLTPGKRHRILIINTSVEN	270
QY		271	SFGLYPADTDAIDRLPEFKVIYASDGLEHAD---TSLYIISAEREVAVDFESDYAGK	327
Dd		271	NFYQSYIGHS-----MIVIESOF-----VPVDSFTLDLSLFEVIGIGQRVDYI IDASQATDN	319
QY		328	TIELRNIGSGSIGIGTDTDYNDNTKVMEKFVAVDDTTPDPTSVVPA-----NL-----	374

RESULT	10			
LAC4_TRAVE				
ID_LAC4_TRAVE	STANDARD:	PRF:	520 AA.	
AC	Q12719;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Laccase 4 precursor (EC 1.10.3.2) (benzenediol:oxyen oxidoreductase)			
DE	(Trisitol oxidase) (Diphenol oxidase).			
GN	LC4 OR LC1.			
OS	Trametes versicolor (White-rot fungus).			
OC	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;			
OC	Ascomycota; Basidiomycota; Coriolaceae; Trametes.			
OX	NCBI_TaxID=5325;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=PRU 572;			
RA	JMELINE=95399443; PubMed=7669813;			
RA	Jonsson L.J., Sjoestrom K., Haegstroem I., Nyman P.O.;			
RT	Characterization of a laccase gene from the white-rot fungus			
RT	Trametes versicolor and structural features of basidiomycete			
RT	laccases. "			
RL	Biochim. Biophys. Acta 1251:210-215(1995).			
CC	-1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED			
CC	PRODUCTS (PROBABLE).			
CC	-1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzosemiquinone + 2			
CC	H(2)O.			
CC	-1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU			
CC	CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE			
CC	3 OR COUPLED BINUCLEAR (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.			
CC	-1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.			
CC	CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	CC			
DR	EMBL; X84683; CAAS9161.1; -			
DR	HSSP; P37064; IAOZ.			
DR	InterPro; IPR001117; Cu-oxidase.			
DR	InterPro; IPR002355; Multicu_oxidase2.			
DR	Pfam; PF000394; Cu-oxidase; 3.			
DR	PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.			
DR	PROSITE; PS00080; MULTICOPPER_OXIDASE2; FALSE_NEG.			
KW	Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;			
KW	Lignin degradation; Multigene family.			
FT	SIGNAL	1	22	POTENTIAL.
FT	CHAIN	23	520	LACCASE 4.
FT	DOMAIN	24	149	PLASTOCYANIN-LIKE 1.
FT	DOMAIN	161	303	PLASTOCYANIN-LIKE 2.
FT	DOMAIN	370	491	PLASTOCYANIN-LIKE 3.
FT	DISULFID	107	509	PROBABLE.

Query Match	7.0%;	Score 215;	DB 1;	Length 520;	Indels 23
Best Local Similarity	21.5%;	Pred. NO. 3.3e-08;			Gaps 4
Matches 119;	Conservative 60;	Mismatches 219;			

01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)

Query Match	6.98;	Score 212;	DB 1;	Length 520;
Best Local Similarity	21.58;	Pred. No. 5.4e-08;		

Matches 119; Conservative 60; Mismatches 219; Indels 156; Gaps 23;

```

OY 34 HEKRYAQSPPYPMETVLPPIPVKQPRLTNTNPVNGEIMWEVEIKPFHQVYDGL 93
DB 15 HSEGRVS-----AALGPVTD--LTISN-----GDVSPDGF 42
OY 94 SADLVGYDGMSPGPTFOVPRGVETVVRFINNAE-----APNSVHLGHSFRAAFDGMAR- 147
DB 43 TAAAVLANGVPPGLITGKGNFQIVINDJLNETMLSTSIHMGFPQKGT--NMAAG 100
OY 148 -----DTEPGSFKDYIYPPNROSARTLWYDHAMHITAENNAKQACLYMLTPDAEDA 200
DB 101 AAFVNOCPATNGNSFL-YDFATDAGTFWYHSH--LSTGYCDGLRQPMVYVDPSPHADL 158
OY 201 LALPSGYGFDIPMLTSSQYANGNL-----VTTNGLNSFWGD-----VIHV 244
DB 159 YV-----DDETTITLSWYHTAASLGAAFLGSDSTLNG--LGRFAGSDSTDLAVITV 212
OY 245 NGQWPKFKNVEPRKYRFRFLDAVSRSF-----GLYFADTDALIDTRLPRFYIASDSGL 297
DB 213 E-----QGRKRYRMLLSISCDPNVYSIDGHNMTIIEADAVN-----249
OY 298 LEHPADTSLIYSMAERYEVF-----DESDYAGKTIELNLGSGIGCTDIDYDNTDKV 353
DB 250 -HEPLTVSIOIYAGQSFVLTAADODIDNFIHALPSAGTTSFSDGINS-----AI 300
OY 354 MRFVVA---DDTTPDQTSVVP---ANLRDVPSPPT-----TNPDRFGRGTGPT 398
DB 301 LKYSAGEVDPTTETTSVLPIDEANL--VPLDSPAAPDDPNIGVDYALNDFNDGNG 358
OY 399 WTINGVAFAADVONRLANVPVGTVERMELIANGN-----GWTN 436
DB 359 FFINDVSEFSPVYLIQILSGTTSADLLPSGSLFAVPSNSTIEISFITAANAPGACH 418
OY 437 PHLHLVDKVSRTSGNNAFTVMPESGLKDYVMLGRETYVEAHYAPFGVYMFCH 496
DB 419 PHLHGTFSIVRTAGSDITFNVPVR---RDVNTGVGVNDVITRETTDNGPMPHLCH 475
OY 497 NLIHEDHDMMAFN 510
DB 476 IDFHLEAGFAIVFS 489

RESULT 12
LACS_TRAVE STANDARD: PRT: 527 AA.
AC 012717:
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Laccase 5 precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)
DE (Urishiol oxidase) (Diphenol oxidase) (Laccase IV).
DE LCC5 OR LCCIV.
OS Trametes versicolor (White-rot fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Coriolaceae; Trametes.
OX NCBI_TaxID=5325;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=527;
RX MEDLINE=97464057; PubMed=9322748;
RA Ong E., Pollock W.B., Smith M.;
RT "Cloning and sequence analysis of two laccase complementary DNAs from
RL the ligninolytic basidiomycete Trametes versicolor.";
RL Gene 196:113-119(1997).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLY).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.

```

-1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.  
 CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL: U44431; AAC49829.1; .
DR HSSP: P37064; IAO2.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; Multicopper oxidase.
DR Pfam: PF00394; Cu-oxidase; 3. Multicopper oxidase.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; FALSE NEG.
DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; FALSE NEG.
KW Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;
KW Lignin degradation; Multigene family.
FT SIGNAL 1 23
FT CHAIN 1 24
FT DOMAIN 25 150 PLASTOCYANIN-LIKE 1.
FT DOMAIN 162 306 PLASTOCYANIN-LIKE 2.
FT DOMAIN 373 498 PLASTOCYANIN-LIKE 3.
FT METAL 87 87 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 89 89 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 132 132 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 134 134 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 425 425 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 428 428 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 430 430 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 480 480 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 481 481 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 482 482 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 486 486 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 486 486 COPPER (TYPE 1) (BY SIMILARITY).
FT CARBOHYD 74 74 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAc. . .) (POTENTIAL).
SQ SEQUENCE 527 AA; 56094 MW; D9597491F1F79825 CRC64;

```

Query Match 6.8%; Score 210; DB 1; Length 527;  
 Best Local Similarity 22.6%; Pred. No. 7.7e-08;  
 Matches 119; Conservative 61; Mismatches 214; Indels 132; Gaps 26;

```

OY 88 VYPD--LGSADLVGYDGMSPGPTFOVPRGVETVVRFINNAE-----APNSVHLGHSFRA 140
DB 38 VPPDGTTRAAVLAG--GVPPGLITGNKDEQIVINDJLNETMLKSTTIHMGIFQ-- 93
OY 141 AFDGMAR-----DTEPGSF-KDYIYPPNROSARTLWYDHAMHITAENNAKQACLY 191
DB 94 ACTINWADGAFAFNOCPIATGNSFLYDFVPOQ--AGTFWYHSH-----LSTGYCDGLRPL 147
OY 192 MLTPDAEDALNLPSCYGEEDIPMLT-----SKQYANGNLVTTNGLNSFWGD 240
DB 148 VYDPPDANASL--YVDVDDTFTVITLADWYHTAAKLGAPAPAGDPSYLING--LGRFSGD 203
OY 241 -----VIHNGQWPKFKNVEPRKYRFRFLDAVSRFGLYFPDQTAIDRLPRFYIA 292
DB 204 GGGATNLVITV-----TQGRKRYRMLLSISCDPNV-----TSSIDGH-----NMTI 245

```



QY 293 SDGSLLEHPA-DTSLIYSAERIEVDFSDYAGKITE-----LRNLGSGIGIGTDT 345  
 Db 246 IEVGVNHEADVDSDIIFAGORSFILN-----ANOSIDNWIRAIPTG-----TTDT 295  
 QY 346 DYDNTDKMREVVADD---TQGPDSVVPANLRD-VPPSPST-----TNTPROFRF 392  
 Db 296 TGVNSALIKRDTAEDEPTTATTSVPLTETDLVPLDNPAAADPOVGVGLAMSLDF 355  
 QY 393 GRTPTWTINGVAFADYONRLANVPVGTVERMELINAGMT----- 435  
 Db 356 SFNSNFFINNETVPPVPLVLIOLSGAODASLLPNGSVYTLPSNSTIEISPIITTD 415  
 QY 436 -----HPHILHLDVKYISFTSGNNATVMPYESGLKDVVWLGRRTVVEAHYAP 486  
 Db 416 GALNAPGAPRPHLGHTEFSVRSAGSSTFNYPVR---RDTVSTGNSGDNVTIRFTTD 472  
 QY 487 FPGYMFHCHNLHEDHMAAFNATVLPDGYNATVFPVPMELM 532  
 Db 473 NPGWFLHCHIDFHLDD---AGFAIVEADTADTAS--ANVPYPAW 512

RESULT 13  
 LACS\_TRAVI STANDARD; PRT; 527 AA.  
 AC 099056;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Laccase 5 precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)  
 DE (urishiol oxidase).  
 GN LCC5.  
 OS Trametes villosa (White-rot fungus).  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 OC Aphyllophorales; Coriolaceae; Trametes.  
 ON NCBI\_TaxID:47662;  
 RX SEQUENCE FROM N.A.  
 RX MEDLINE:97128774; PubMed:8973314;  
 RA Yaver D.S., Gollightly E.J.;  
 RT Cloning and characterization of three laccase genes from the  
 RT white-rot basidiomycete Trametes villosa: genomic organization of the  
 RT laccase gene family.?  
 RT Gene 181:95-102(1996).  
 RN [2]  
 RP REVISIONS.  
 RA Yaver D.S., Gollightly E.J.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED  
 CC PRODUCTS (PROBABLE).  
 CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2  
 CC H(2)O.  
 CC -1- COFACTOR: BINS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU  
 CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE  
 CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.  
 CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: L78078; AAB47735.2; -  
 CC HSSP: P37064; IAOZ.  
 CC InterPro: IPR001149; Cu-oxidase.  
 CC InterPro: IPR002355; Multicopper oxidase.  
 CC Pfam: PF00394; Cu-oxidase; 3.  
 CC PROSITE: PS00079; MULTICOPPER\_OXIDASE1; 1.

DR PROSITE: PS00080; MULTICOPPER\_OXIDASE2; FALSE NEG.  
 KW Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;  
 KW Lignin degradation; Multigene family.  
 FT SIGNAL 1 23  
 FT CHAIN 24 527  
 FT DOMAIN 25 150  
 FT DOMAIN 162 306  
 FT DOMAIN 373 498  
 FT METAL 87  
 FT METAL 89  
 FT METAL 132  
 FT METAL 134  
 FT METAL 145  
 FT METAL 425  
 FT METAL 428  
 FT METAL 430  
 FT METAL 480  
 FT METAL 481  
 FT METAL 482  
 FT METAL 486  
 FT METAL 486  
 FT CARBOHYD 74  
 FT CARBOHYD 77  
 FT CARBOHYD 156  
 FT CARBOHYD 209  
 FT CARBOHYD 233  
 FT CARBOHYD 242  
 FT CARBOHYD 276  
 FT CARBOHYD 317  
 FT CARBOHYD 358  
 FT CARBOHYD 366  
 FT CARBOHYD 393  
 FT CARBOHYD 402  
 FT CARBOHYD 464  
 FT CARBOHYD 464  
 SQ SEQUENCE 527 AA; 36247 MW; FIB638D65FEA478 CRC64;

Query Match 6.7%; Score 207; DB 1; Length 527;  
 Best local Similarity 22.9%; Pred. No. 1.3e-07;

Matches 113; Conservative 56; Mismatches 200; Indels 124; Gaps 24;

QY 88 VYPD--LGSADLVGYDMSPPPTQVPRGYEVYVFINNAE-----APNSVHLHGSFSA 140  
 Db 38 VFPDGLTRAAVLAVG--GVFPGLITGKKGDEFOINVINDLNTNMLKSTIIHMHGIFQ-- 93  
 QY 141 AFDGMAE-----DTEBGSF-KDYIYNNROSARKLWYHDAHMTTANAYRGAGLY 191  
 Db 94 AGTNMADGAAFVNCQPIATGNSFLYDFTVDDO--AGTFWYHSH---LSTOYCDGLRGPL 147  
 QY 192 MLTDPADALNLPSCGCEFDIPMLT-----SKQYTANGNLVYTNGLNSFWCP 240  
 Db 148 VYDPPDDPNASL---IDVDDDTVITLADMYHTAAKIGAPFAGPDSVLING--LGRSGD 203  
 QY 241 -----VIHVGQWPFKNVPRKRYREFELDAVSRSGLYFADTDAIDTRLPKRYA 292  
 Db 204 GGGATNLVTIVT-----TQGRKRYRFLVSLSDPN-----TFSLDGH-NMTIIE 247  
 QY 293 SDGSLLEHPADTSLIYSAERIEVDFSDYAGKITE-----LRNLGSGIGIGTDT 346  
 Db 248 VD--GVNHEALVDSDIIFAGORSFILN-----ANOSIDNWIRAIPTG-----TTDT 296  
 QY 347 YDNTDKMREVVADD---TQGPDSVVPANLRD-VPPSPST-----TNTPROFRF 393  
 Db 297 TGVNSALIKRDTAEDEPTTATTSVPLTETDLVPLDNPAAADPOVGVGLAMSLDF 356  
 QY 394 RTGPTWTINGVAFADYONRLANVPVGTVERMELINAGMT----- 435  
 Db 357 SFNSNFFINNETVPPVPLVLIOLSGAODASLLPNGSVYTLPSNSTIEISPIITTDG 416  
 QY 436 -----HPHILHLDVKYISFTSGNNATVMPYESGLKDVVWLGRRTVVEAHYAP 487  
 Db 417 VLNAPGAPRPHLGHTEFSVRSAGSSTFNYPVR---RDTVSTGNSGDNVTIRFTTD 473  
 QY 488 FPGYMFHCHNLH 500

Db 474 PGWFLCHIDFH 486

RESULT 14

LAC2\_CORHI STANDARD: PRT: 520 AA.

ID LAC2\_CORHI

AC 002497;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Laccase precursor (EC 1.10.3.2) (benzenediol: oxygen oxidoreductase)

DE (urushiol oxidase) (ligninolytic phenoloxidase).

OS Coriolus hirsutus.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Coriolaceae; Coriolus.

OX NCBI\_Taxid=5327;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN-IFO 4917;

RA MEDLINE-90368706; PubMed-2394718;

RA Kojima Y., Tsukuda Y., Kawai Y., Tsukamoto A., Suglura J.,

RA Sakano M., Kita Y.;

RT "Cloning, sequence analysis, and expression of ligninolytic phenoloxidase genes of the white-rot basidiomycete Coriolus hirsutus."

RT J. Biol. Chem. 265:15224-15230(1990).

CC -1- FUNCTION: MOST PROBABLY PLAYS AN IMPORTANT ROLE IN LIGNIN DEGRADATION. CLEAVES THE C-C AND C-O BONDS OF SOME PHENOLIC LIGNIN MODEL COMPOUNDS (SUCH AS O- AND P-QUINOLS, AMINOPHENOLS AND PHENYLENEDIAMINE).

CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) -> 4 benzoemiquinone + 2 H(2)O.

CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- POLYMORPHISM: 2 ALLELIC FORMS VARYING IN ONE AA POSITION.

CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.

CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: M60560; AAA33103.1; -

DR EMBL: M60561; AAA33104.1; -

DR PIR: A35883; A35883.

DR InterPro: IPR001117; Cu-oxidase.

DR InterPro: IPR002355; Multicu\_oxidase2.

DR Pfam: PF00394; Cu-oxidase; 3

DR PROSITE: PS00079; MULTICOPPER\_OXIDASE1; 1.

DR PROSITE: PS00080; MULTICOPPER\_OXIDASE2; FALSE\_NEG.

KW Oxidoreductase; Signal; Copper; Metal-binding; Lignin degradation;

KW Glycoprotein; Repeat.

FT SIGNAL: 1 21

FT CHAIN: 22 520

FT DOMAIN: 23 148

FT DOMAIN: 160 302

FT DOMAIN: 369 491

FT METAL: 85 85

FT METAL: 87 87

FT METAL: 130 130

FT METAL: 132 132

FT METAL: 416 416

FT METAL: 419 419

FT METAL: 421 421

FT METAL: 473 473

FT METAL: 474 474

POTENTIAL.

LACCASE.

PLASTOCYANIN-LIKE 1.

PLASTOCYANIN-LIKE 2.

PLASTOCYANIN-LIKE 3.

COPPER (TYPE 2) (BY SIMILARITY).

COPPER (TYPE 2) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 1) (BY SIMILARITY).

COPPER (TYPE 2) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 1) (BY SIMILARITY).

FT METAL 475 475

FT METAL 479 479

FT METAL ?

FT CARBOHYD 50 50

FT CARBOHYD 72 72

FT CARBOHYD 75 75

FT CARBOHYD 210 210

FT CARBOHYD 229 229

FT CARBOHYD 354 354

FT CARBOHYD 457 457

FT DISULFID 106 509

FT DISULFID 138 226

FT VARIANT 411 411

FT CONFLICT 378 379

SQ SEQUENCE 520 AA; 55688 MW; 977D8DFA551F7929 CRC64;

Query Match 6.7%; Score 205.5; DB 1; Length 520;

Best Local Similarity 23.0%; Pred. No. 1.6e-07; Indels 137; Gaps 29;

Matches 121; Conservative 67; Mismatches 202;

QY 87 QVYPDIGSADLVGDMSPGPTFQVPGVETVVFRTINAE-----APNSVHLGSEFSRAA 141

DB 35 EVSPDGFARQAVVNNVTPEGLVAGNKGDFOLNVDNLNHTMLKSTSHMHGFFOKG 94

QY 142 FDCMAE-----DITEGSEF-KDYIYRNKRSARLWTHDHAMHTTAENARYGQAGLYW 192

DB 95 --WMADGPAFVNOCPISSGHSFLYDFQVDPQ--AGTFWYHSH---LSTGYCDGLGPFV 146

QY 193 LHPDAEDALNLPSCGSEFDPMLTFSKOYTA-----NG---MLVTNGE 233

DB 147 VYPPNDPRLASLYVDND-DVITLADWYHTAAKLGAFPLGADATLNLGGRSPSTTAAD 205

QY 234 LNSFWGSDVHVNQGPWFKNVEPRKYRFRFLDAAVSRSGFLGADDAIDFLRPFVIAS 293

DB 206 L-----AVINV-----TKGRYRFRFLVSL-----SCDPNHTFSIDGH 237

QY 294 DSGLLP-----HPDTSLLYISMAERYEVVFPFSDYACKTIELR---NLG--GSIGGIC 342

DB 238 DLITIEVDSINSQPLVYVDSIQIFPAQRYSPVLNADQVQ--NYWIRANPNPENGWVAGGIN 296

QY 343 TD-TDYNPDKVFARFVADDDTOPDTSVVPANLRD-----VPPSP---TNTPRGF 390

DB 297 SALIRTDGADPV-----EPTTQTTPTRKPLNEVDLHPLATMAVP--GSPVAGVDALINM 349

QY 391 RFGRTSPRTWINGVARADYQONRLLANVPQTVVERWELINAGN----- 432

DB 350 AFGNPGNPFINGASFVPPVPLQLIISGAQNAQDLPLSGSVSLPSNADIEISPPATA 409

QY 433 ---GWHPIHILVDFKVLISRTSGNNARVMPYESGL-KDYVNLG---RRETVVVEAHYA 485

DB 410 AARGAPRPFELHGAFAVY--RSAGS---TVYNYNDPIFRDVSSTGTPAADNTIRFR-T 464

QY 486 PFGVYMFHCHNLIHEDHOMMAFNAFVLPDYGNATVFPVDPMEELM 532

DB 465 DNGPWFELHCHIDFLEAGFAVVF-AEDIDP-----VASANFVPOAM 505

RESULT 15

LAC2\_TRAVE STANDARD: PRT: 519 AA.

ID LAC2\_TRAVE

AC 012718; 012718;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Laccase 2 precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)

DE (urushiol oxidase) (diphenol oxidase) (Laccase I).

GN LCC2 OR LCC1.

OS Trametes versicolor (White-rot fungus).

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Coriolaceae; Trametes.

OX NCBI\_Taxid=5325;

RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-523;  
RX MEDLINE-97464057; PubMed=9322748;  
RA Ong E., Pollock W.B., Smith M.;  
RT "Cloning and sequence analysis of two laccase complementary DNAs from  
the lichenolitic basidiomycete *Trametes versicolor*.";  
RL Gene 196:113-119(1997).  
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED  
PRODUCTS (PROBABLE).  
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2  
H(2)O.  
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU  
CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE  
3 OR COUPLED BINUCLEAR (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.  
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U44851; AAA86659.1; -  
DR EMBL: U44430; AAC9828.1; -  
DR HSSP: P37064; IAOZ.  
DR InterPro: IPR001117; Cu-oxidase.  
DR InterPro: IPR002355; MultiCu\_oxidase2.  
DR Pfam: PF00394; Cu-oxidase; 3.  
DR PROSITE: PS00079; MULTICOPPER OXIDASE1; 1.  
DR PROSITE: PS00080; MULTICOPPER OXIDASE2; FALSE\_NEG.  
KM Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;  
Lignin degradation; Multigene family.  
FT SIGNAL 1 20 POTENTIAL.  
FT CHAIN 21 519 LACCASE 2.  
FT DOMAIN 22 147 PLASTOCYANIN-LIKE 1.  
FT DOMAIN 159 301 PLASTOCYANIN-LIKE 2.  
FT DOMAIN 368 490 PLASTOCYANIN-LIKE 3.  
FT METAL 84 84 COPPER (TYPE 2) (BY SIMILARITY)  
FT METAL 86 86 COPPER (TYPE 3) (BY SIMILARITY)  
FT METAL 129 129 COPPER (TYPE 3) (BY SIMILARITY)  
FT METAL 131 131 COPPER (TYPE 3) (BY SIMILARITY)  
FT METAL 415 415 COPPER (TYPE 1) (BY SIMILARITY)  
FT METAL 418 418 COPPER (TYPE 2) (BY SIMILARITY)  
FT METAL 420 420 COPPER (TYPE 3) (BY SIMILARITY)  
FT METAL 472 472 COPPER (TYPE 3) (BY SIMILARITY)  
FT METAL 473 473 COPPER (TYPE 1) (BY SIMILARITY)  
FT METAL 474 474 COPPER (TYPE 3) (BY SIMILARITY)  
FT METAL 478 478 COPPER (TYPE 1) (BY SIMILARITY)  
FT METAL ? ? COPPER (TYPE 1) (BY SIMILARITY)  
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CONFLICT 69 69 D -> V (IN REF. 1; AAC49828).  
SQ SEQUENCE 519 AA; 55810 MW; BFFB5B4CD0007702 CRC64;

Query Match 6.7%; Score 205; DB 1; Length 519;  
Best Local Similarity 23.7%; Pred. No. 1.7e-07;  
Matches 124; Conservative 63; Mismatches 203; Indels 134; Gaps 30;

QY 88 VPPDGSADLVGCTDGMSPPTQVGRGVTVVRFNNAE-----APNSVHLGSPS 138  
DB 35 VSPDGLRDALIVNGVPSPLITGRKG---DRFOLNVDDTLNHSMLKSTSIHWHGFQ 90

QY 139 RAAFDGMAE-----DITEPGSF-KDYVYYPNROSARTLWYHDHAMHITAENAYRQAG 189  
DB 91 -AGTNWADGAPFVAVNOCPLASGHSFLYDFHPVQ--AGFTWYHSH-----LSTQYCCGLNG 142  
QY 190 LYLWTDPAEDALNLPFGYGEEDIP-----MILTSKOYTANGNLVTNGELNSFWGDIYHV 244  
DB 143 PFVYVDPKPBHAS-----RYVDNNESTVITLTDWYHTA-----ARLGRFFPLGADATLIL 191  
QY 245 NG-----QPMFKNVE-PKRYFRFLDAVNSFGILYFADTALDTRLPFKYIAD 294  
DB 192 NGIGRSASTPTAALAVINVOHGKRYRFLVLSICDPNY-----TFSDIGH-NLFTVIEVD 244  
QY 295 SGLLEPADTSLIYISMAERYEVFPFSDYAGK--TIELRNIG--GSIGIGTDTDYDMT 350  
DB 245 -GINSQPLVVDLSIQITAAQRTSFVLANOTGVNMYWRANPNSTGVFAGINS----- 296  
QY 351 DKVMRF--VVADDTQOPDTSVVP--ANLRDVP-PPSPPTNTP-----RQFRFGRTGP 397  
DB 297 -ALIRYQAPVAPVAPPTTQTSVIFLETMLHPLARMVPGSPTPGGVDKALNLAFNENG 355  
QY 398 TWITINGVADVOGNRLLANVPVGTVERWELIANGN-----GWTN 436  
DB 356 NEFINNATPTPTVPVLLQILSGAQTAODLLPAGSVYPLPAHSTIBITLPATYALAPGAP 415  
QY 437 PIHILVDEKVISRTSGNNARFVMPYESGL-KDVVWLG--RRETVVEAHVAPFGVYV 492  
DB 416 PRLHGHAFAYV-RSAGS---TTYNTNDPIFRDVSSTGTPAAGDNTIR-FQTDNFGPMF 470  
QY 493 FPHCNLIHEDHDMMAAFNATVLPDVGYNATVFVDPMEEELMQARP 536  
DB 471 LRCHIDFILD---AGF-----AIVFAEDVADYKAAMP 499

Search completed: October 3, 2002, 11:02:18  
Job time: 442 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 11:03:21 ; Search time 55.98 seconds

(without alignments)  
1767.651 Million cell updates/sec

Title: US-09-656-640A-4

Perfect score: 3067

Sequence: 1 MFKHTLGAALSLIFNSNAV.....AVTERIQTMARYRYAADE 572

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPREMBL\_19:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-virus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1913	62.4	602	09P8C3	09P8C3 acremonium
2	653.5	21.3	475	093M03	093M03 streptomyc
3	506	16.5	568	10	023123 arabidopsis
4	502	16.4	637	10	09FTS3 oryza sativ
5	495	16.1	582	10	09AMU4
6	485	15.8	581	10	0949X9
7	470	15.3	591	10	09C9A4
8	456	14.9	591	10	09FTS5
9	454.5	14.8	614	10	09FTS6
10	400.5	13.1	527	16	067206
11	400	13.0	536	2	0938E6
12	393	12.8	515	16	09CJ06
13	384.5	12.5	502	16	098F88
14	288	9.4	468	16	09CP81
15	276.5	9.0	1662	2	P71431
16	276	9.0	513	16	09PME8

17	266.5	8.7	460	2	09X3V2	09X3V2 pseudomonas
18	261	8.5	463	16	09HXW7	09HXW7 pseudomonas
19	260	8.5	500	16	09K849	09K849 bacillus ha
20	242.5	7.9	520	3	059944	059944 ceriporiops
21	241.5	7.9	504	16	053858	053858 mycobacteri
22	229.5	7.5	518	3	059896	059896 pycnoporus
23	229	7.5	464	16	092QV6	092QV6 rhizobium m
24	225.5	7.4	518	3	09UV02	09UV02 pycnoporus
25	225.5	7.4	561	3	096WNO	096WNO dolitylus cl
26	224.5	7.3	520	3	013448	013448 coriolus ve
27	223.5	7.3	520	3	096UT7	096UT7 trameles ve
28	222.5	7.3	518	3	09HDS9	09HDS9 polyporus c
29	221.5	7.2	518	3	096FR6	096FR6 pycnoporus
30	220.5	7.2	518	3	096V45	096V45 pycnoporus
31	220	7.2	539	3	09Y780	09Y780 coprinus cl
32	220	7.2	567	10	0941X2	0941X2 oryza sativ
33	220	7.2	569	10	09FY79	09FY79 arabidopsis
34	217.5	7.1	519	3	013420	013420 basidiomyc
35	216	7.0	352	2	093F57	093F57 pseudomonas
36	215.5	7.0	520	3	094222	094222 trameles ve
37	215.5	7.0	524	3	013422	013422 basidiomyc
38	215.5	7.0	589	3	09C497	09C497 glomerella
39	215	7.0	352	2	093F48	093F48 pseudomonas
40	215	7.0	586	10	09AU16	09AU16 pinus taeda
41	214.5	7.0	351	2	093F47	093F47 pseudomonas
42	214.5	7.0	526	3	013456	013456 coriolus ve
43	214.5	7.0	577	10	09FUD5	09FUD5 arabidopsis
44	213.5	7.0	353	2	093F51	093F51 pseudomonas
45	213.5	7.0	533	3	060199	060199 pleurotus o

## ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	602 AA.
09P8C3	09P8C3			
AC	09P8C3:			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	POLYPHENOL OXIDASE PRECURSOR.			
GN	PPOA.			
OS	Acromonium murorum.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;			
OC	Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Acromonium.			
OX	NCBI_TaxID=45278;			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=CBS 157.72;			
RX	MEDLINE=21268855; PubMed=11375170;			
RA	Gouka R.J., van der Heiden M., Swarthoff T., Verrips C.T.;			
RT	"Cloning of a phenol oxidase gene from Acromonium murorum and its			
RT	expression in Aspergillus awamori."			
RL	Appl. Environ. Microbiol. 67:2610-2616(2001).			
DR	EMBL: AJ271104; CAB75422.1; -			
KW	Signal.			
FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	62	602	POLYPHENOL OXIDASE.
SQ	SEQUENCE	602 AA;	66920 MW;	0303D991A05228A3 CRC64;
Query Match		62.4%;	Score 1913;	DB 3; Length 602;
Best Local Similarity		61.0%;	Pred. No. 2.7e-135;	
Matches 365;		Conservative 73;	Mismatches 126;	Indels 34; Gaps 8;
QY	5 TLAAALSLIFNSNAVQASP-----VPE-----TSPATHLFKRY 39			
DB	4 TTRLRALALYLSKGAQAMKFEIDPEEFAALAAIIVEDDPANDIORSPADN--LQRR 61			
QY	40 AQTSPQPMFTVLPPIPKQPRLTVTNPNGOEIYVEVEIKPFTHQVYPDGSADLVG 99			
DB	62 SPSPAVYTLFGQAPLPSIPVKEPELFTVTNPNGEIDYIEIEIKHFSQVFPDGLPADLVG 121			

Oy	100	YKDSPGGTFOVPGREVEVVFVFINNAEPNSVHLHGSSRAFPQGMEDITEPESFDY	129
Oy	122	YDGISPGGTLOEKGRESEVVFVFKATVESSILHLGSTRAPWQGMEDVITNGETVDY	161
Oy	160	YPNROSARTLWYHDBAHMHTAEMARYRQAGLMLTDPDAEDALNDPSGGEFDIPMLTISK	219
Db	182	YPNAGAGFSWYHDBAHMHTAEMARYNQAAGYIIHDDPAEDSLGPSYGEYDIPLLTISK	241
Oy	220	OYTANGLVTTNGELANSWGVYIHVNGQPMFVKVYERPKTRFRFLDAVARSGLYFADT	279
Db	242	QYNSDGLTFESTKGEQSLMGVIOVNGVPMPEYFVEYERPKYRFRLLDAVARSFSLYVDT	301
Oy	280	DAIDRLPEFKYIASDGLLEHPADTSLIYIMAREYEVFPSPDACKTIELRLGSGIS	339
Db	302	ADEDTIRIFQVIYASDGLLEEPYTTSKLYSIARREYELILDSPSEKTIELRN-EPAGV	360
Oy	340	GIGTDTDYDNTDKWRFVYAD-DTTPDTSVVPANLRDVPESPPTNT-PROFNRGTGP	397
Db	361	GLGIEVNYDDDTDKWRFVYAGVGLSSPDTSVSPSTLRDVPESPSTITIDHSFPFATAG	420
Oy	398	TWTINGVAFADVQNRLLANVGVGVYERKRELINAGNGWTHPIHLVLPFKYISRI--SGN	454
Db	421	QMSINGVYFSVDEKRLLANVPLGTVOIMQLTNAAQGWTHPIHLVLPFKILSRFGATGA	480
Oy	455	NARTWMPYS-GLKDVYVGLGRETVVEAHYAPPPGYMPCNHLIHEDHDMMAAFNATV	513
Db	481	TTTRGVPEPSAGLNDVYILGGEVYVYEAHAPPPGYMPCNHLIHEDHDMMAAFNATV	540
Oy	514	LPDYGYNATVADVPMEELMQRAPYELGFEQASQGSFVSQAUTERTIQTMAEYRPPYAAD	571
Db	541	LPDYGYNSTALADPDEFFRAKPYVDNDVEVERNAFSTDEIEAQVLMASLYLPYDND	598
RESULT	2		
ID	Q93M03	PRELIMINARY;	PRT; 475 AA.
AC	Q93M03;		
DT	01-DEC-2001 (TREMblrel, 19, Created)		
DT	01-DEC-2001 (TREMblrel, 19, Last sequence update)		
DT	01-DEC-2001 (TREMblrel, 19, Last annotation update)		
DE	OXIDOREDUCTASE-LIKE PROTEIN.		
GN	AUR21.		
OS	Streptomyces aureofaciens.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;		
OC	Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomycetes.		
OX	NCBI_TaxID=1894;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CCM3239;		
RA	Korhaneec J., Bistakova J., Novakova R., Homeroova D., Rezuchova B.,		
RT	"Cloning and characterization of a new polyketide gene cluster in		
RL	Streptomyces aureofaciens CCM3239."		
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY033994; AAK61713.1;		
DR	SEQUENCE 475 AA; 52602 MW; A2ED9A4C638DDOC9 CRC64;		

	Query Match Similarity	21.3%	Score 653.5	DB 2:	Length 475:	
	Best Local Similarity	36.1%	Pred. No. 7.1e+11:			
	Matches 174:	Conservative	60:	Mismatches 189:	Indels 55:	Gaps 19
QY	40	AQISQYEMFTYPLRIPP-VKQPRLVTPNPVNGQELIYVEVEIKPTTHOVYPDLGSADLV	98			
Db	31	AEAAIVP-FANAMPLPKRLKPTSPFATISDL-----YEIRNGQQLVYVGLWS-KVR	81			
QY	99	GIDGMSPGCFTEQVPRGVETVTRF-INNAEAPNSVHLHGSFSRAAFDGMADITPEGSFKDY	150			
Db	82	TYDGTFPEPTIRATQGREVVYRQINELQVNTAVHLGCAVLSSEHGLPMDITVPGGERTY	141			
QY	159	YYPNQGSRATLVYHHNHAMHITAEENYRQGAGLYMLTRDPEDALNLPSGGEEDIPILITS	218			
Db	142	RYPNQGPASATLVYHHNHAEENYFMGLHGLYLLTDHNEKRLPLPS--GPDVAVPLVID	199			

Qy	219	KOYANGSLVTTNGBELNFMGDVH--VNGCPMEFKVVERPKYRFRLLDAVASSECEYE	276
Db	200	ARVEDGTLTRP-----SDCPHMLNGKEREYFOVAAKRTERYNNACANNKYKLR	253
Qy	277	ADTDADTRLEPKVIASDGLLEHPADTSLIYISMAEREYVFPDSOY-AGKTIELRLNG	335
Db	254	AD-----GIEFTQIGDGGFLLEPRVOOSELLMIGGERADIVDFSKRYKGDVSYLENPG	307
Qy	336	G-SIGIGCTDIDYDNTDKVMRFRVADDTTOPDTSVVPANLRDVEPSPNTNTPROFRGR	394
Db	308	AQSI-----ERPEVMRFDIV--RTADYSFVFGRLTYV-PQPTLYVERDFEI-R	353
Qy	335	TGPTTINGVAFADYQNFLLANVEYGIVERKELIN---AGNG-----WTHPIHLIYVD	444
Db	354	TEFPATTINGOSYD--PNKVIDTARLGTTEVTVYRNVEAPARACGRDFEHLN-HSEHTEILTY	410
Qy	445	EKVISRTSGNNAARYVMPEESLJKDVMILGRETLYVVEANHARFPGVYMFCHNLIHEDHD	504
Db	411	FRVLER---NGKPRAGTRDLKDKPTVLLGPDGTIVKIAMTWGPRYGYOLHYCHDGHSSGG	466
Qy	505	KM	506
Db	467	QM	468

RESULT	3	
023123		
ID	023123	PRELIMINARY; PRT; 568 AA.
AC	023123;	
DT	01-JAN-1998 (TREMBLrel_05, Created)	
DT	01-JAN-1998 (TREMBLrel_05, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel_19, Last annotation update)	
DE	FL19G10.5 PROTEIN.	
CN	FL19G10.5.	
OS	Arabidopsis thaliana (Mouse-ear cross).	
OC	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsis.	
OX	NCBI_TaxID=3702;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CV, COLUMBIA;	
RA	Federstepf N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,	
RA	Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,	
RA	Oji O., Osborne B.I., Shlun P., Sun H., Toriumi M., Vytotskaia V.,	
RA	Yu G., Ecker J., Theologis A., Davis R.W.;	
RL	Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AF000657; AAB72167.1; to the EMBL/GenBank/DBJ databases.	
DR	InterPro: IPR001117; Cu-oxidase.	
DR	Pfam: PF00394; Cu-oxidase. 1.	
QO	SEQUENCE 568 AA; 64355 MW; C4314C889576E35E CRC64;	

		16.5%;	Score 506;	DB 10;	Length 568;	
		Best Local Similarity	29.3%;	Pred. No. 1,1e-29;		
		Matches 168;	Conservative	70;	Mismatches 180;	Indels 156; Gaps 27
Oy	48	MTETVLPPIPRVKQPLRTVTNPVNGOEIWIYEVEIKPFTQV-----YPDLSGADLV	98			
			:	:	:	:
Dd	29	MEVDLLP-----DMPRLFGNSVHG-----IIKPASLQIGFSPKMKFHNDLPATPVF	76			
			:	:	:	:
Oy	99	CYDGMS-----PGTFQVPRGCVLEVVRIN-----NAEAPNSVH	132			
			:	:	:	:
Dd	77	AV-GFSRSKAIVPGPTIEIVGVDTVVYVNRHLPKSHILPWDPPTSIPATPKKGIGITVVH	135			
			:	:	:	:
Oy	133	LHGFSRAAFDGMADIEDTPEGSGFKD-----YYUPRQSARTLYMDHAMHIATAENA	183			
			:	:	:	:
Dd	136	LHGTHHEPTSDONA-DAMWTAGFRETRGPKWTKTTTLHYEKKQORGNMNYIDHAMGLLRVYL	194			
			:	:	:	:
Oy	134	YRGAGLIYMLDPA--EDALNLPSGSGEEDIIMILTSKOYTANG---NIVLTNGELINSFW	238			
			:	:	:	:
Dd	195	LAGLVGAAILRHNAVAESRPOLPTG--DEFRPRLIIDPRSRRKGSUYMMNTGGNNPSHPOM	253			
			:	:	:	:

```

OY 229 ---GDVIVHGQWPEPKNVPEPKYRFFELIAAASRSGLEAFADDAIDTRLPKEVIAS 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 254 OPEYFGVDIIVNGKAMPILNRRRRKRIREFIINASNAERKFFFSN-----GIDFIVGS 307
OY 294 DSGLEHREADTSILATISMAERYEVVEFSDYAGKTLELNGLSGIGIGIDTDY----- 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 308 DSAVLSKRVPMRKILLSLSPSEILVDVVVDFKSPSRIVLAN-----DAPYRPPGSD 357
OY 348 ---DNTDKVMRVVADDTTQPTSVYRANLRVPPPSPTTNPPOFRGFRGTGTWINGV 404
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 PVNENGKVMRKII--NNESEDCTCPKYL--INPNADVSAVLRVY-----ISMYEV 409
OY 405 AFADQNRLLAN-----VPV-----GTVERMELINAGGNWHPHIIHIVDEKVISRTS- 452
    : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 SNSDEPTHLVNGLEPYEARVETETPKRSGTEVEVNLNLED--NPLHILGLFRVKEQAL 468
OY 453 -----GNNA-----FTVMEYESGLRDVWL--GRETVEVEE 483
Db 469 LAAGLEEKECMTKQNDAAVKQISKYARGKTAAYNAHEGKMNVFMFMGHYTRILVRS 528
OY 484 YA-----PF-----PCVYMFHCHLILHEDHDM 506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 YIHTNASYPEDPTQBPQ--YVYHCHILHEDHNNMM 561

RESULT 4
Q9FTS3 PRELIMINARY; PRT; 637 AA.
AC Q9FTS3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE SPORE COAT PROTEIN-LIKE PROTEIN.
GN P0409B08.14.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eubatidoideae; Oryzaeae; Oryza.
CX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT clone:p0409B08.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002860; BABJ8287.1; -.
DR InterPro; IPR002355; MultiCu_oxidse2.
DR PROSITE; PS00080; MGLTICPPER_OXIDAS2; UNKNOWN_1.
KW Coat protein.
SQ
SEQUENCE 637 AA: 70919 MW: 692D39F10F2D6A96 CRC64:

```

	Query Match	16.4%	Score 502:	DB 10:	Length 637:	
	Best Local Similarity	29.1%:	Pred. No. 2.6e-29:			
	Matches	174:	Conservative	72:	Mismatches 186:	Indels 164:
	Gaps					29
Oy	24 PVPEPSPATG-----HLEK-----RVAQISPOYPMETVPLIPPVKKOPRLTYVN	67				
Dd	82 PLP--APAAGRMRMVMNLSLNTLKTKYVDLSLRITAKINGVGIRGHAPVPDI-----KTTI--	131				
Oy	68 PVNG--OELWYEVEIKPFTHQGVYDGLSGADLVGDGM-----PGPTFOYPRGGTENV	119				
Dd	132 ---GMSTKWQFHRMPRPFPVFY-----GSSLQTATFRPGPIIVARQGPHNAV	176				
Oy	120 RFINN-----AEAPNSVHLHGSEFSRAAFDMA-----EDLTSPSS	154				
Dd	177 EQMNHLPDANILPMDDPKVTALPRKKGQVPTYVNLHGCAHPREDHGNAFMFTRDFAENSS	236				
Oy	155 F---KDYLYPNQSAKTLMTYNDHANHTAEMNYRGQAAGLMYLTDP-AEALNLPSCGE	210				
Dd	237 TWTRKYTVYPNQADAGNLTMDHNAGLTFVSLLAGLLAAYILEKEPLEDEPNMLPC--GGH	294				

QY	211	DIMILTSKQYTAGNLTVTGELNSFEGDVIYHNQGPWPPEKNTPEPKKYAFRFLDAVSR	270
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	295	DLHLVIADREYVTGSISSIDREMKPEYFGLTVITNGKAMPPLSVQRRRLRILINASMAR	354
QY	271	SFGLYFADTDIDLRLPEKVASDSGLLEHNPADTSLLYISMAEKYEVWPFPSDVAKTIE	330
		: : : : :	
Db	355	YFNWTL-----NGALPPTVYIGSSSYLSRPVYTSNLYLSAELEFDYIVDESRLPAMTE	409
QY	331	LRNIGS-----IGGIGTDYDNTDKVMEFVA-----DDTTPQDTSVVPANLRDPF-	379
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	410	TEMLTADYPRPENGPNVTDPLDGG-KYMLFEKVAGKKQVD--MPDKSKYBEH--GVPA	463
QY	360	-----PSTTTPPROFGRGRTGP--TWITNGVAFADVQNRLLANPV-----GTVEWEL	427
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	464	SVALPRLPTTTRYLYLVENQYAPRNLYINGRLLED-----PYMETPEKSGTTELMQV	514
QY	428	INAGGWHPIRHILHVDKVSRT-----SGNNARPTVEESGLDKDVMYLGK	474
		: : : : :    : : : : :    : : : : :    : : : : :	
Db	515	INL-TCGNHPLHLHATFOALTKMKIEGFQVEFKDCMKNNNTATCUNLDQHAVGVVAPPE	573
QY	475	RETV-----VEAHA-PP-----PCVVMHCHNLIDHEHDMM	506
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	574	EKTKNAVKATLRPEMTSVVAVAFRLVENQYPRPDATTEEG-EYHGHILDHEDNMAM	630

ID	Q9AMU4	PRELIMINARY:	PTT:	582 AA.
AC	Q9AMU4			
DT	01-JUN-2001	(TREMBLrel. 17, Created)		
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	PURATIVE SPORE COAT PROTEIN.			
GN	P0044F08.18 OR P0037C04.31.			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehharitoidae; Oryzaeae; Oryza.			
OX	NCBI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. NIPPONBARE;			
RA	Sasaki T., Matsumoto T., Yamamoto K.;			
RT	"Oryza sativa nippobare(Ga3) genomic DNA, chromosome 1, PAC			
RL	clone:P0044F08.";			
RP	Submitted (NOV-2000) to the EMBL/genbank/DBJ databases.			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. NIPPONBARE;			
RA	Sasaki T., Matsumoto T., Yamamoto K.;			
RT	"Oryza sativa nippobare(Ga3) genomic DNA, chromosome 1, PAC			
RL	clone:P0037C04.";			
RL	Submitted (FEB-2001) to the EMBL/genbank/DBJ databases.			
DR	EMBL, AP002903; BAB5542.1; -			
DR	EMBL, AP003233; BAB5542.1; -			
DR	InterPro: IPR002355; MultiCu.Oxidase2.			
DR	PROSITE: PS00080; MULTICOOPER_OXIDASE2; UNKNOWN_1.			
KW	Coat protein.			
SC	SEQUENCE 582 AA; 64050 MW; B5054BECBAC5CIDA CRC64;			
QY	Query Match	16.1%;	Score 495;	DB 10; Length 582;
QY	Best Local Similarity	29.2%;	Pred. No. 7.6e-29;	
QY	Matches 175; Conservative	73;	Mismatches 219;	Indels 132; Gaps 27
QY	9 AALSLLEFNSNAQAQSPVETSPATGHLERVA-----QISPOXYM---PLP	54		
QY	8 AVLLLVIVGVAGCTRRPSAPRPVETDLQKAVGSLSEMYDELQPMKTIQFGSMRNCHNSP	67		
QY	55 IPRVQPRLLTVTNPNVG--QETWYELVELIKRTHQYVPRDLGSAADLVGVGMSRPRTFOVP	112		
QY	68 I-----RLTI-----GMVQKKMKHRRDLPASTVVEFGTSAA-----ATPPGPTIEAA	110		

Qy	113	RGVEYVREFINNAE-----	-PMSVHG-----	-SSRAAFQSMAE	147
Db	111	QGVPLSTWQYTLPRRIILPMDPPVPAIPRRGVPTVVIHGGAAHPQSDGSFAMFETA	170		
Qy	148	DITEPG---SEKDYVYPPNRSQARTLAWHDHAMHTAENAVRGQAGLYMLTPADEA	203		
Db	171	GGEGEGPAMSTPTTYTPYPAQSGVGLWHDHMLGLTRANLLAGLLGAVVINPVEAPLGL	230		
Qy	204	PSGIGEDPIPIMLTSKQTYANGNL-VTTNGELNS-----	EMGDYIHNQGPWPFXNV	254	
Db	231	PCG-DEFQRLVLMADRSPYADGSIYMYNTGIIPIPHQMPQEPYGEAITYNGKAMPFLAV	289		
Qy	255	EPKRYRFFFLAAVRSGLFYAFDADTDIDLTPKVLVASDGLLEHPADSLIYISMAER	314		
Db	290	ARRRRRPIIITNSARNTNL-----SITNLPPTVYVGGSDTNYLSKRYTAAASLLVSAET	343		
Qy	315	YEYVFEFSDYAGKTIELRNIGSGSIGIGTDTDXTDKVMKFEVYVADDTGTPDTSVPANL	374		
Db	344	FDVVVDFSGTSSEAEELVNTAPRYPDG-QAPNDLNGKVMKFEVISPAPAK-DTSRPVAKL	401		
Qy	375	RDVEPSPSTMTPR-----QFRGRG-PIWT-INGVAFADQYRLLANVPQYVERWE	426		
Db	402	LDYVAVAEEBAVQRYIYMEYTEDAATGNPHLYXNGKRLDEPATE--TPRQTEYWE	458		
Qy	427	LINAGMGTHPIHILHLYDFKV-----ISRTSGNNAATVM	460		
Db	459	VINLTPD-NHRLHLHLATFQATRVRGVLVDEDAFKGCAKMLNDARCVSHNAVGEBAVAP	517		
Qy	461	PYESGLKQVYVL--GRRETVVYEAHYA-----PF-----PGVMFHCHMLIHEDHDM	506		
Db	518	EHKGMKNVYKIAAGYMTIYVKEFFMYDSGKRPYDPAETAEFG-VYVCHILIDHEDNMI	575		

Query Match	15.88;	Score 485;	DB 10;	Length 581;
Best Local Similarity	30.28;	Pred. No. 4.2e-28;		
Matches 156;	Conservative 66;	Mismatches 183;	Indels 112;	Gaps 23

```

0y 91 DLGASDLVGYDCMS-----PGPTPQVPRGVETVVRFIN-----N 124
    | | : : | | | | | | : | | : |
Db 84 DLPRATPVFAI-GTSKRSATVPRPTTEAVGVVDVYVTVRNHLPLNHLPLWDPRTISPAIPKH 142

```

```

0y 125 AEPASVILHOSFSRAAFDGAED-----ITEPS---FDDYYVKNROSARTLMYDHAM 176
Db 143 GGIPTVYHLHGHCHIEPPSDGNADSWFAGREKIOSKTKTKTTHVYNKQOPGNMYHDHAA 202
0y 177 HITABENAYRGQAGLXMLT-DPAEDALNLPSCYGEEDIPMLTISKQYANG-----NLVTN 231
Db 203 GLTRNNLGLAGLGGYILRHSSVESBPLPLPNG-REFDEPLVIFPBRSEFKSDSIYMNATGN 261
0y 232 GELNSFW-----GVIVHNGGQPMFPFKVVEPRKYFRFLDAVRSSEFLYPADDAIDTRL 286
Db 262 PTIHQWQPEYFGDAILYNGKAMPRLVRRKRKYRTITMSNARFFPEFFSN-----GL 315
0y 287 PPKYIASDSCGLEHPADTSLYIISMAERYVEDFSDYAGKTIELRNLGSGIGETDID 346
Db 316 DFIWVSGDSAYLARPVSTRKSYLLAPSEIVLVLDVFSKSTKATLANNADPYPSGDPYT 375
0y 347 YDNTDKVRFVVAADDTQOPDTSVPANILRDVPFSPPTNPROFR-----GRGTPT 398
Db 376 EENS-KVMKFLI-NYKSDVDSIILPKKL-IEYPAHVSTSTRTRYIAMEEYVSIDEPT 431
0y 399 WY-INGCAFADYONRLANPY-VCTVERWELINAGNCWTHPIHILYDFV-----447
Db 432 HLYINGLPY---NAPYTEPKTIDTSWEWEINLTED-NHRPLHILGLFKVLEBOTALVNS 486
0y 448 -----ISRTSGNNAFTVMPVESEGLKDYVWL--GRRETVVVEAHY-----484
Db 487 EEFTECMKRNDAVKCELSKARGENKTAIVYVHEGGMKNVFKMMPGHYTKILVAFSYHSN 546
0y 485 -----APFGVYMFCHNLIEEDHDMMAAFNATVL 514
Db 547 ESYSEFDAQTEPG-VYVCHCHIDHEDNNMMRPF-AMVL 581

```

RESULT	7		
Q9C9A4		PRELIMINARY;	PRT; 591 AA.
Q9C9A4			
Q9C9A4			
AC	01-JUN-2001 (TREMBLrel, 17, Created)		
DT	01-JUN-2001 (TREMBLrel, 17, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel, 17, Last annotation update)		
DE	SPORE COAT PROTEIN-LIKE PROTEIN.		
GN	F23N20.3.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eumariota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:		
OC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:		
OC	eurosid II: Brassicales: Brassicaceae: Arabidopsis.		
OX	NCBI_TaxId=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV, COLUMBIA.		
RX	MEDLINE=21016719; PubMed=11130712;		
RA	Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,		
RA	White O., Alonso J., Altai H., Arujo R., Bowman C.L., Brooks S.Y.,		
RA	Buehlerr E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,		
RA	Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,		
RA	Dunn P., Etgu P., Feilblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,		
RA	Gill J.E., Goldsmith A.D., Johns B., Hansen N.F., Hughes B., Huizar L.		
RA	Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,		
RA	Kim C.-J., Koo H.L., Kremetskala I., Kurtz D.B., Kwan A.I., Lam B.,		
RA	Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,		
RA	Lih X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,		
RA	Miltscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,		
RA	Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,		
RA	Sakano H., Salerg S.L., Schwartz J.R., Shinn P., Southwick A.M.,		
RA	Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,		
RA	Underback T., Van Aken S., Vaysberg M., Vysotskala V.S., Walker M.,		
RA	Wu D., Yu G., Fraser C.M., Wenter J.C., Davis R.W.,		
RT	"Sequence and analysis of chromosome 1 of the plant Arabidopsis		
RT	thaliana."		
RL	Nature 408:816-820(2000).		
DR	EMBL: AC016972; AAC51692.1; .		
Q0	Coat protein.		
Q0	SEQUENCE 591 AA; 67358 MW; B0CE7B0E8F05AFA CRC64;		





RT clone: P0044F08.  
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AP002860; BAB18284.1; -  
 DR EMBL: AP002909; BAB21194.1; -  
 DR InterPro: IPR002355; MulticOxIdase2.  
 DR PROSITE: PS00080; MULTICOOPER\_OXIDASE2; UNKNOWN\_1.  
 KW Coat protein.  
 SO SEQUENCE 588 AA; 64290 MW; 6C6B995060C4A82 CRC64;

Query Match 14.8%; Score 454.5; DB 10; Length 588;  
 Best Local Similarity 26.6%; Pred. No. 8.4e-26;  
 Matches 163; Conservative 78; Mismatches 189; Indels 183; Gaps 28;

OY 9 AALSLFFNSNAVQASVPETSPATGHLEFRVAQISQVPMFVPLPIPPVQKPRILVTNP 68  
 DB 37 ADLQVAGSLQMTYVADLPMAKIRGIFGORGAV-----P 71  
 OY 69 VN-----GQELIYVEVEIKP---FTHQVYPDLSADLVGYDMSPGPTQVPRGVETVVR 120  
 DB 72 ILTLTGWQKTKKFRHDLPTATVFVYGGCPD-----SATFPQPTIMARHDPVLEFVR 122  
 OY 121 FVN-----NAEPNSVHLGS-----FSRAAFDGMADIIEPG-- 153  
 DB 123 WENHLPSHILPMDPTVPATPAIKNGVPTVVAHLHGSAPHDSDGSFAFMTAGFAEKGA 182  
 OY 154 -SEKQVYVYVNSASRTLMYHDHAMHTTAENAVRGAGLMTLDPADA-LNLPSCYGEFD 211  
 DB 183 WTQATYRPRVNPVPGNLMYHDHALGLTRANLLAGLGAIVIEKPEVDYMDLPC--DDDD 240  
 OY 212 IPMLITSKOYANGNLVYTNGLNS-----FMGDVIVHNSQMPPEKKNVEBR 257  
 DB 241 LHLVLAIDRSFNVDSGLY-----MNSTGVAPNIIHPQWAPPEFGEALIVNGKAMPFLVYHRR 295  
 OY 258 KRFREFLDAAVRSRGVYFADDAIDTRLPRKYIASDSGLLEHPRATSLIYXMAERYEV 317  
 DB 296 RYRLRLNLSNARYFNV-----SLNSGLPIHVVGSDASYLSAPVYSNLLSPATIFIV 349  
 OY 318 VDFSDYAGKTIELRN-----LGSGIGIGTDEYDIDTKVMRFVAVDD--TTQPDTS 368  
 DB 350 VDFGOSPTAELELNSAARYPTTGAAGPRL-----NGVMKRIYVQNPGRDLPDPS 401  
 OY 369 VVPAALRDVPPS-----PTNTPTQPRFGRTGPTWT-----INGVAADVONRLIANP 418  
 DB 402 TVPDH--EVPYASVATLPTPTMTRYIVMEYELTPGOSTHLYINGLRLD-----P 450  
 OY 419 V-----GVEWELLINAGNGTHPIHILVDKVIS-----RTSGNNA-- 456  
 DB 451 VETPKSGTTELMQYINL--TGDNHPLHLHGMFOAKMOQLVNLQAFDTCMTAVINDAKC 509  
 OY 457 -----RTVMP--YESGLKDVVWL--GRRETVVVEAHYA-----PGVYMF 493  
 DB 510 NVDQNAVGVEVVPDHEKTKWKVIVKPRGFVYSVIAFLVDTNQTYPPTDTAERG--YYX 568  
 OY 494 HCHNLIHEDHDM 506  
 DB 569 HCHILHEDNAMI 581  
 RESULT 10  
 ID 067206 PRELIMINARY: PRT: 527 AA.  
 AC 067206;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)  
 DE PERIPLASMIC CELL DIVISION PROTEIN (SUFI).  
 DE SUFI OR AQ\_1130.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 OX NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus".  
 RT Nature 392:353-358(1998).  
 DR EMBL: AE000724; AAC07157.1; -  
 DR InterPro: IPR001117; Cu-oxidase.  
 DR InterPro: IPR002355; MulticOxIdase2.  
 DR Pfam: PF00394; Cu-oxidase; 1.  
 DR PROSITE: PS00080; MULTICOOPER\_OXIDASE2; 1.  
 KW Cell division; Complete proteome.  
 SO SEQUENCE 527 AA; 59490 MW; F361BF791EF97355 CRC64;

Query Match 13.1%; Score 400.5; DB 16; Length 527;  
 Best Local Similarity 30.1%; Pred. No. 8.2e-22;  
 Matches 146; Conservative 72; Mismatches 202; Indels 65; Gaps 22;

OY 70 NGQELIYVEVEIKPFTHOYVPLDLSADLVGYD-GMSPGPTQVPRGVETVVRFINNAEAP 128  
 DB 59 DQQRV---STAKMTTLEVIIPG-KSTDMLVYIDNEYVIFLRKGQTFSSADPVNNSGED 114  
 OY 129 NSVHLHGSFSAFADG---MAEDITEPGSFKDYVYVNSASRTLMYHDHAMHTTAENAVR 185  
 DB 115 STIHHMGFRAPKSDGHPYAVAKDDETSYPRFTIIDRSG--TYFHPHPRGRTGYQVY 172  
 OY 186 GQGLYMLTDPADALN--LPSCYGEFDIPMLITSKOYANGNLVYTNGLNSFMGDVYH 243  
 DB 173 GLAGMIIIEDEDEDLKQALDLEYGVIDIPLIIDOKTFPSSQQLVYVPMGHMGFWGDTLL 232  
 OY 244 VNGQPEFKNVEPRKRFELDAVRSRGVYFADDAI---DTRLPRKYIASDSGLLEH 300  
 DB 233 VNLTPHYDVERKIRYRFTILNGSNARPRRL-----ALLRGNQRRFVIVIEGGLDT 286  
 OY 301 PADTSLIYSMAERYEVDFSDYA-GKTIELRN-----LG-GSIG-----G 340  
 DB 287 PREVNIIIVAPGERIDILVDFDASVNDVIKLYNPHNLIGMIGMRGMGMEROMG 346  
 OY 341 IGTDDY---DNTD-KVMFVVAADTTPDTSVPAALRDVPPST-----TNTPR 388  
 DB 347 NGMNMDMGNADNSEEFEVMEFRVTKDSAY-DKS-IPQRLSEVT-PIVTDGAQVQRTTLGMR 403  
 OY 389 OFRFGRTGPTWINGVAFADVON-RLANVPVGTVERMELINAGNGTHPIHILVDK 447  
 DB 404 RAVFTINGETWEDGYANPDINNPKYLFEQNNGDVYILEYV--NIGMTHPIHILHGFQ 462  
 OY 448 ISRTSGNNARTVMPYESGLKDVVWLGRRETV--VVEAHYAPFPGVYMFCHNLIHEDH 504  
 DB 463 LRRSLGRLRAT---DLGKQIVYIAPMETVRIADVMSHRYNHNQYLLHCHILEHDEG 518  
 OY 505 MMAAP 509  
 DB 519 MMVNY 523  
 RESULT 11  
 ID 0938E6 PRELIMINARY: PRT: 536 AA.  
 AC 0938E6;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
 DE MULTICOOPER OXIDASE.  
 DE CUID.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.



QY 2 EKHTLGAALSLLENSNAVQASPVPE TSPATGHLFKRVAQISQYPMFTVPLIPPVKQP 64

SO SEQUENCE 1662 AA; 174293 MW; 639EE23860DD9246 CRC64;

Query Match 9.0%; Score 276.5; DB 2; Length 1662;  
Best Local Similarity 19.9%; Pred. No. 8,7e-12;  
Matches 144; Conservative 59; Mismatches 165; Indels 357; Gaps 26;

```
OY 49 FTVPLPDP-----PVKQRLVTNPNVNGQEIW--YYEVEIKPETHQVDP 91
      |||:|
      |||:|
DB 101 FIDPLPLPGKLSLADGVTRTIPVAVPAKWI-NPQTGAATSDDYFEIALVEYKOKLHSD 159
      |||:|
      |||:|
OY 92 L-GSADLVGY-----DG-----102
      |||:|
      |||:|
DB 160 LKNPTTLRGVYDLSVDGTGLYYPNSSSGKVAPEANADAPADAKPVMINAVDAKGFMTGA 219
      |||:|
      |||:|
OY 103 -----MSP--GPTFOYPRGVEYVVRFIN-----NAEAP-----128
      |||:|
      |||:|
DB 220 KYQARVYMERPMYGLIQARKGPTPLKTFYNLLPGRAETTVGADGKVQVYARNGDIFLP 279
      |||:|
      |||:|
OY 129 -----NSVHLHG-----135
      |||:|
      |||:|
DB 280 LDKSIHAGLGPGFTEFTQNRNSHILHGDDTPWISDGTPHQWITPIEENANANPKALVN 339
      |||:|
      |||:|
OY 136 -----SFSKAAPDGAEDITEPGS--FKDYIYPNROSARTLWYHDAMHTTAENAYR 185
      |||:|
      |||:|
DB 340 QGIDPEFLPSFLRGASQWYPMDPDGAGASTYTFPNGOSARMLWYHDHTIGVTRLNVYA 399
      |||:|
      |||:|
OY 186 GOAGLYMLTDPAD-----ALNLPSCYGEFDIPMITSKOYT-----ANGNLY 228
      |||:|
      |||:|
DB 400 GMAAYTTLDEVDLTKTGTGALKVLPAPADTIPLVLTDRTFVPADVALODARWNTS 459
      |||:|
      |||:|
OY 229 TTNGE-----LNSF-----237
      |||:|
      |||:|
DB 460 AMGESDSWFPHYETVQDPNNGNFSYGRMHMGFWPVPFAMYDLPSGEXGDVTVTP 519
      |||:|
      |||:|
OY 238 --WGDYIHNQGPWPKNPERKRYRFLDAVRSKF--GLYFAD-----278
      |||:|
      |||:|
DB 520 EAMMDPPLVNGVAYPTIELDPKVRMKVLNASNDRFNISLFAVDAEQRINDPLLGATE 579
      |||:|
      |||:|
OY 279 ---TDAIDTRLP-----FKVIASDS 295
      |||:|
      |||:|
DB 580 VKMVDAAVSATFCACAGVTRAVVATDGSYCTPETWPTDNRPGVPSPAAGPSFFQIANEG 639
      |||:|
      |||:|
OY 296 GLEHPAD-----TSLYISMAERYEVVDFSDYAGKTIELRNL 334
      |||:|
      |||:|
DB 640 GLLPKVAELAPTPVGYQLDKGRITVNLVLTGLYLGNABERADVLVLSAVAGKTLIVND 699
      |||:|
      |||:|
OY 335 GGS-----IGIGTDTDY-----DNTDKVMRFVYVADDTQPDTSVPA 372
      |||:|
      |||:|
DB 700 SGAPVPACDPRNDYFTAVGDOSDAGCAEDTKPGYGNTRTMQIKVRAAITPSS--FDG 756
      |||:|
      |||:|
OY 373 NLRDVPFPSPPT---NTPRQRFGRGTGPTWITNGVAFADYONRLLANPVGIVERWELI 428
      |||:|
      |||:|
DB 757 QIRDAQRQSDSTALKETKATKAYAIAGEKP-----VVGODVYNOAL---GTT--WGAT 803
      |||:|
      |||:|
OY 429 NAGNG 433
      |||:|
      |||:|
DB 804 PSLNG 808
```

Search completed: October 3, 2002, 11:03:23  
Job time: 462 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:52 ; Search time 25.3 Seconds  
(without alignments)  
552.231 Million cell updates/sec

Title: US-09-656-640A-4

Perfect score: 3067

Sequence: 1 MFKHTLGAALSLFLNSNAV.....AVTERIQTMAYRYPAADE 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/3A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCrUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfillsl.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3067	100.0	572	US-09-401-476-4	Sequence 4, App1
2	2010	65.5	583	US-09-401-476-2	Sequence 2, App1
3	228.5	7.5	499	US-09-032-315-2	Sequence 2, App1
4	228.5	7.5	499	US-08-993-318A-2	Sequence 2, App1
5	228.5	7.5	499	US-09-399-886-2	Sequence 2, App1
6	228.5	7.5	499	US-09-396-260-2	Sequence 2, App1
7	228.5	7.5	499	US-09-576-281-2	Sequence 2, App1
8	228.5	7.5	520	US-08-462-484-2	Sequence 2, App1
9	228.5	7.5	520	US-08-441-147-2	Sequence 2, App1
10	228.5	7.5	520	PCr-US95-07536-2	Sequence 2, App1
11	226	7.4	620	US-08-940-661A-2	Sequence 2, App1
12	226	7.4	620	US-09-083-485-2	Sequence 2, App1
13	226	7.4	620	US-08-939-218A-2	Sequence 2, App1
14	223.5	7.3	511	US-08-462-484-8	Sequence 8, App1
15	223.5	7.3	511	US-08-441-147-8	Sequence 8, App1
16	223.5	7.3	511	PCr-US95-07536-8	Sequence 8, App1
17	220	7.2	539	US-09-032-315-1	Sequence 1, App1
18	220	7.2	539	US-08-993-318A-1	Sequence 1, App1
19	220	7.2	539	US-08-689-421-27	Sequence 27, App1
20	220	7.2	539	US-09-399-886-1	Sequence 1, App1
21	220	7.2	539	US-09-396-260-1	Sequence 1, App1
22	220	7.2	539	US-09-399-528-27	Sequence 27, App1
23	220	7.2	539	US-09-181-827A-27	Sequence 27, App1
24	220	7.2	539	US-09-576-281-1	Sequence 1, App1
25	215.5	7.0	616	US-08-991-531-2	Sequence 1, App1
26	215.5	7.0	616	US-09-032-315-9	Sequence 9, App1
27	215.5	7.0	616	US-08-993-318A-9	Sequence 9, App1

28	215.5	7.0	616	US-09-028-887-2	Sequence 2, App1
29	215.5	7.0	616	US-09-399-886-9	Sequence 9, App1
30	215.5	7.0	616	US-09-396-260-9	Sequence 9, App1
31	215.5	7.0	616	US-09-518-901-2	Sequence 2, App1
32	215.5	7.0	616	US-09-576-281-9	Sequence 9, App1
33	214.5	7.0	616	US-08-749-882A-2	Sequence 2, App1
34	214.5	7.0	616	US-08-539-134-2	Sequence 2, App1
35	212.5	6.9	616	PCr-US95-06816-2	Sequence 2, App1
36	211.5	6.9	516	US-08-689-421-29	Sequence 29, App1
37	211.5	6.9	516	US-09-389-528-29	Sequence 29, App1
38	211.5	6.9	516	US-09-181-827A-29	Sequence 29, App1
39	210.5	6.9	529	US-08-172-931B-14	Sequence 14, App1
40	210.5	6.9	529	US-08-706-037-25	Sequence 25, App1
41	210.5	6.9	529	US-09-005-397-25	Sequence 25, App1
42	210.5	6.9	529	US-09-032-315-5	Sequence 5, App1
43	210.5	6.9	529	US-08-993-318A-5	Sequence 5, App1
44	210.5	6.9	529	US-09-399-886-5	Sequence 5, App1
45	210.5	6.9	529	US-09-396-260-5	Sequence 5, App1

#### ALIGNMENTS

RESULT 1  
US-09-401-476-4  
; Sequence 4, Application US/09401476  
; Patent No. 6168936  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Huaming  
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes  
; FILE REFERENCE: GC584  
; CURRENT APPLICATION NUMBER: US/09/401.476  
; CURRENT FILING DATE: 1999-09-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Stachybotrys charatum  
US-09-401-476-4

Query Match	100.0%	Score 3067	DB 4	Length 572
Best Local Similarity	100.0%	Pred. No. 2e-289		
Matches 572	Conservative 0	Mismatches 0	Indels 0	Gaps 0
OY	1	MFKHTLGAALSLFLNSNAVQASVPETSPATGHLFRVAQISPOYPMFTVPLPIPVKQ	60	
Db	1	MFKHTLGAALSLFLNSNAVQASVPETSPATGHLFRVAQISPOYPMFTVPLPIPVKQ	60	
OY	61	PLRTVTPVNGQELWYVEVEIKPPTHQVYVYDLSADLVGDSMSPGTFQVPRGVETVVR	120	
Db	61	PLRTVTPVNGQELWYVEVEIKPPTHQVYVYDLSADLVGDSMSPGTFQVPRGVETVVR	120	
OY	121	FINNAEPNVSVHLHGSFRAAFDGMADITEPGSFQYYPNROSARTLWYHDAHMITA	180	
Db	121	FINNAEPNVSVHLHGSFRAAFDGMADITEPGSFQYYPNROSARTLWYHDAHMITA	180	
OY	181	ENMYRGOAGLYMLTDPEDALNLPSPGYGFDPIMILTSKQYANGNLVTNGELNSWGD	240	
Db	181	ENMYRGOAGLYMLTDPEDALNLPSPGYGFDPIMILTSKQYANGNLVTNGELNSWGD	240	
OY	241	VIHVNGQPMWPKKVPKRYFRFLDAAVSSFGLYFADTDAIDRLPFKYASDSGLLEH	300	
Db	241	VIHVNGQPMWPKKVPKRYFRFLDAAVSSFGLYFADTDAIDRLPFKYASDSGLLEH	300	
OY	301	PADSLIYISMAERYEVEFSDVAGKTIELRNIGSISGIGTDVDDNDKVMRFVAD	360	
Db	301	PADSLIYISMAERYEVEFSDVAGKTIELRNIGSISGIGTDVDDNDKVMRFVAD	360	
OY	361	DTTQOPTSVVPAANLROVPFSPPTTNTPRQFRGCTGTWTINGVAFADVONRLLANVPV	420	
Db	361	DTTQOPTSVVPAANLROVPFSPPTTNTPRQFRGCTGTWTINGVAFADVONRLLANVPV	420	

QY 421 TVERMELINANGWTHPHIHLVDFKVISRTSGNNARTVWPYESGLKDVYMLGRRETVV 480  
DB 421 TVERMELINANGWTHPHIHLVDFKVISRTSGNNARTVWPYESGLKDVYMLGRRETVV 480  
QY 481 EAHYAFPPGYVMHCHNLHEDHDMMAAFNATVLPDYGATVFPVDMELMOARPELG 540  
DB 481 EAHYAFPPGYVMHCHNLHEDHDMMAAFNATVLPDYGATVFPVDMELMOARPELG 540  
QY 541 EFOAGSGFVQAVTERIOTMAERYPAADE 572  
DB 541 EFOAGSGFVQAVTERIOTMAERYPAADE 572

RESULT 2  
US-09-401-476-2  
; Sequence 2, Application US/09401476  
; Patent No. 6168936  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Huang  
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes  
; FILE REFERENCE: GC584  
; CURRENT APPLICATION NUMBER: US/09/401,476  
; CURRENT FILING DATE: 1999-09-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 583  
; TYPE: PRT  
; ORGANISM: Stachybotrys chartarum  
US-09-401-476-2

Query Match 65.5%; Score 2010; DB 4; Length 583;  
Best Local Similarity 65.4%; Pred. No. 1,1e-186;  
Matches 383; Conservative 71; Mismatches 110; Indels 22; Gaps 10;

QY 1 MKRHTGAAALSL-LFNSNAVQASPPV--ETSPATGHLKRVQAIPOVPMFTVPLPIP 57  
DB 1 MTSQAIGAVLGLAVIGSSVDARSVAGRSTDMPSG-LTKROTLSPLALVEVPLPIP 59  
QY 58 VKOPRLTVNPVNGOEIWEVEIKPPTHQVYRDLGSADLVGYDGMSPPTPOVPRGVT 117  
DB 60 LKAPN-TVPMPNTGEDILYEMELRPSHOIYDLEPANNVGYDGMSPGPTIIVPKTES 118  
QY 118 VYRFINNAE--APNSVHLGHSFSAFDDGAEEDITEPGSEKDYYPNROSARTLWYHDA 175  
DB 119 VYRFVNGENTSPNSVHLGHSFSAFDDGAEEDITEPGSEKDYYPNROSARTLWYHDA 178  
QY 176 MHTAENAVRGAGLYMLDPAEDALNLSGVEEDIPMLITSKOYTANGNLVTTNGELN 235  
DB 179 MSTAEANAVMGAGVYIMIDDPADALNLSGVEEDIPMLITSKOYTANGNLVTTNGELN 238  
QY 236 SFGADVHNNGOWPKNVEPRKYRFRFLDAAYSRSGLYEADTDALDTPKRVASDS 295  
DB 239 STWGDVIYQNGOWPKNVEPRKYRFRFLDAAYSRSGLYEADTDALDTPKRVASDS 298  
QY 296 GLEHHPADTSLIYISMAEREVEVDFSDYAGKTIELNLGSGIGICTDIDYDNTDKVMR 355  
DB 299 GLEGGVDDITDLYISMAEREVEVDFSDYAGKTIELNLGSGIGICTDIDYDNTDKVMR 357  
QY 356 EYVADDTTPODTSVDPANLKDVPF-----SPTTNPPOFRRGSGTGTWTINGVAA 407  
DB 358 EYVDEVLESPTDSEVPANLKDVPFPEGGMNDPANPTDD--ETFTFGRANGOWTINGVTF 415  
QY 408 DVONRLLANVPQSTVERMELINANGWTHPHIHLVDFKVISRTSGNNARTVWPY- 466  
DB 416 DVENRLLRNVPQSTVERMELINANGWTHPHIHLVDFKVISRTSGNNARTVWPY- 472  
QY 467 KDVYMLGRRETVVVEAHYAFPPGYVMHCHNLHEDHDMMAAFNATVLPDYGATVFPV 526  
DB 473 KDVYMLGRRETVVVEAHYAFPPGYVMHCHNLHEDHDMMAAFNATVLPDYGATVFPV 532

QY 527 PMELMOARPELGEFQAGSGFVQAVTERIOTMAERYPAADE 572  
DB 533 PMELMOARPELGEFQAGSGFVQAVTERIOTMAERYPAADE 578

RESULT 3  
US-09-032-315-2  
; Sequence 2, Application US/09032315  
; Patent No. 5985818  
; GENERAL INFORMATION:  
; APPLICANT: Svendsen, Allan  
; TITLE OF INVENTION: LACCASE MUTANTS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5985818 No. 5985818disk of No. 5985818th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032,315  
; FILING DATE: 27-FEB-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rozek, Carol  
; REGISTRATION NUMBER: 36,993  
; REFERENCE/DOCKET NUMBER: 5200,200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 499 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-032-315-2

Query Match 7.5%; Score 228.5; DB 2; Length 499;  
Best Local Similarity 22.8%; Pred. No. 1,1e-13;  
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

QY 55 IPEVKOPRLTVNPVNGOEIWEVEIKPPTHQVYRDLGSADLVGYDGMSPPTPOVPRG 114  
DB 2 IGPVAD-LITMAA-----VSPDGSFRQAVVYVNGGPGGLITGNMG 41  
QY 115 VETVVRFINNAE--APNSVHLGHSFSAFDDGAE-----DITEPGSF-KDYVY 160  
DB 42 DRQLVNDILNHTLMKSTSIHMHGFFQGT--NMAADGPAFINQCPISGSHSLYDFQV 99  
QY 161 PNRQSARTLWYHDAHHTAENAVRGAGLYMLDPAEDALNLSGVEEDIPMLITSKQ 220  
DB 100 PDQ--AGTWYHSH-LSTOYCDGLRPFVYVYDNDPAADLYVDN-----DDTVITLVDM 151  
QY 221 YTANGMLVTTNGELNSFWGVDVHNGO-----PWPEKRYEP-RKYRFRFLDAVSR 270  
DB 152 YHVAAL-----GPAFLGADATILNCKGRSPSTTTADLSVISTPCKRFRFLVLS 207  
QY 271 SE-----GLYPADDAIDTRLPKRVIASDSGLEHHPADTSLIYISMAEREVEVDFSD 323  
DB 208 NYFFSIDGHNMTIIEEDSINTA-----PLVDSIQIFAQORYSFVLE-AN 251  
QY 324 YAGKTIELR--NLG--SGISGIGICTDIDYDNTDKVMR--VVADDTTPODTSVDPANL 375



Db 252 QAVDNYWIRANPNFGVGTGINS-----ALLRYDGA-AAVEPTTQTSTAPLNEY 303  
Qy 376 DVPEPSPTNTPR-----OPFRGRTGPTWTINGVAFADYONRLANVPVGVER 424  
Db 304 NL-HPLVTTAVPGSPVAGVDLAINMAFNNGNFINGASFTPPVPLQIISGAONA 362  
Qy 425 WELINAGN-----GWHPIHILVDFKVISRTSGNNARTVWPYE 463  
Db 363 QDLPSGSVYSLPSNADIEISFPATAAPGAPHPHGHAFVAV--RSAGS---TYNYND 418  
Qy 464 SGL-KDYVWLG---RRETVVEAHYAPFGVYMFHCHNLHEDHDMAAFNATVLPDYGY 519  
Db 419 NPFRDVSSTGTPPAAGDNTVIRER-TDNPQPMFLHCHIDHLEAGFAVVF-AEDIDP--- 473  
Qy 520 NATVEVDPMEEIM 532  
Db 474 --VASANPVQAM 484

RESULT 4  
US-08-993-318A-2  
; Sequence 2, Application US/08993318A  
; Patent No. 5998353  
; GENERAL INFORMATION:  
; APPLICANT: Pedersen, Anders  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Schneider, Palle  
; APPLICANT: Rasmussen, Grethe  
; APPLICANT: Cherry, Joel  
; TITLE OF INVENTION: LACCASE MUTANTS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 59983530 No. 5998353disk of No. 5998353th America  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/993,318A  
; FILING DATE: December 18, 1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gregg, Valera A.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 5032.200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; TELE: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 499 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-993-318A-2

Query Match 7.5%; Score 228.5; DB 2; Length 499;  
Best Local Similarity 22.8%; Pred. No. 1.1e-13;  
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

Qy 55 IPPVQOPRLTVNPNVNGGSIWYVEIKRFTHQVYPDLSADLYGYDGMSPGPTPQVPRG 114  
Db 2 IGPVAD-LITINAA-----VSPDGSROAVVYVNGGPGGLITGNMG 41  
Qy 115 VETVAFINNAE-----APNSVHLHGSFSRAAFDGMAE-----DITEPGSF-KDYVY 160

Db 42 DRFLVYIDNLFTNHTMLKSTSIHMHGFFQKGT--NNADCPAFINQCPISGSHFLYDFQV 99  
Qy 161 PNRQSARTLMLHHHAAHITENAVRGAGLYMLTDPAEDLNLPSGGEFDIMILTSKQ 220  
Db 100 PDO--AGTEWYHSH-LSTQYCDGLRGFEVYDPNDPADLYDVDN-----DQVITLVDM 151  
Qy 221 YTAGNGLVLTNGELNSFGDVIHVNCQ-----PWPKNVEP-RKYRFRFLDAVSR 270  
Db 152 YHVAKL-----GPAFLGADATLNLNGKGRSPSTTADLSVISTPCKRYRFLVSLSDP 207  
Qy 271 SF-----GLYPADDAIDTRLPFKVIASDSGLLEHPADTSLLYISMAERYEVDFSD 323  
Db 208 NYTFSIDGHNMFTIETDSINTA-----PLVDSIQIFAAQRYSEVLE-AN 251  
Qy 324 YAKGTIELR---NLG---GSIGIGITDIDYNTDKVAF---VYADDTQDTSVSPANLR 375  
Db 252 QAVDNYWIRANPNFGVGTGINS-----ALLRYDGA-AAVEPTTQTSTAPLNEY 303  
Qy 376 DVPEPSPTNTPR-----OPFRGRTGPTWTINGVAFADYONRLANVPVGVER 424  
Db 304 NL-HPLVTTAVPGSPVAGVDLAINMAFNNGNFINGASFTPPVPLQIISGAONA 362  
Qy 425 WELINAGN-----GWHPIHILVDFKVISRTSGNNARTVWPYE 463  
Db 363 QDLPSGSVYSLPSNADIEISFPATAAPGAPHPHGHAFVAV--RSAGS---TYNYND 418  
Qy 464 SGL-KDYVWLG---RRETVVEAHYAPFGVYMFHCHNLHEDHDMAAFNATVLPDYGY 519  
Db 419 NPFRDVSSTGTPPAAGDNTVIRER-TDNPQPMFLHCHIDHLEAGFAVVF-AEDIDP--- 473  
Qy 520 NATVEVDPMEEIM 532  
Db 474 --VASANPVQAM 484

RESULT 5  
US-09-399-886-2  
; Sequence 2, Application US/09399886  
; Patent No. 6140092  
; GENERAL INFORMATION:  
; APPLICANT: Pedersen, Anders  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Schneider, Palle  
; APPLICANT: Rasmussen, Grethe  
; APPLICANT: Cherry, Joel  
; TITLE OF INVENTION: LACCASE MUTANTS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 61400920 No. 6140092disk of No. 6140092th America  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/399,886  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/993,318  
; FILING DATE: December 18, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gregg, Valera A.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 5032.200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655

TELEX:  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 499 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-09-399-886-2

Query Match 7.5%; Score 228.5; DB 4; Length 499;  
Best Local Similarity 22.8%; Pred. No. 1,1e-13;  
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

QY 55 IPPVQOPRLVTNPNVNGEIMWEYVEIKPFTHQVYPDLGSADLVGSDGSPGPTFOVPRG 114  
DB 2 IGPVAD--LTITNAA-----VSPDGFSSRAVYVNGGTPGLITGNMG 41  
QY 115 VETVYRFINNAE-----APNSVHLGFSRAAFDGMAE-----DITEPGSF-KDYXX 160  
DB 42 DRFOLNVIDNLNHTMLKSTSIHMGFFOKGT--NMADGPAPINOCPISSGSHFLYDFQV 99  
QY 161 PNRQARFLMTHDHAMHTAENAVRGAGLYMLTPPADALMLPSGYGDFDIPMLITSQ 220  
DB 100 PDQ--AGTFWYHSH--LSTQYCDGLGPFVYVDPNDPADLYVDN---DDTVITLVDM 151  
QY 221 YTANGNLVTTNGELNSFWGDVYIHVNGQ-----PMPFKNVEP-RKYRFRFLDAVSR 270  
DB 152 YHVAAKL-----GPAEPLGADATLNGKGRSPSTTADLSVISVTGKRRYRFLVSLSCP 207  
QY 271 SF-----GLYFADTDALDTRLPKRVIASDSGLLEHPADTSLIYISMAERYEVVDFSD 323  
DB 208 NYTFESIDGHNMTIETDSINTA-----PLVDSIOIFAQORYSFYLE-AN 251  
QY 324 YAGKTIELR--NLG--GSIGIGTDTVDNTDKVMRF--VVAODTQPDTSVVPANLR 375  
DB 222 QAVDNYWIRANPNFGVNGFTGINS-----ALIRYDGAALVEPTTQTSTAPLANEV 303  
QY 376 DVPPSPPTTNTPR-----OFRGRTGPTWTINGVAFADYONRLANVPVGTVER 424  
DB 304 NL-HPLVTTAVGSPVAGVDLAIMMAFNENGTNFINGASTPTPVYLLDIIISGAQNA 362  
QY 425 WELINAGN-----GWTPIHILHVDKVISRTSGNNAKRYMPE 463  
DB 363 QDLPSGSVYSLPSNADIEISFPATAAAGAPHPHILHGAFAVY--RSAGS---TVYND 418  
QY 464 SGL-KDYVWLG---RREYVVEAHYAAPFGVYMFCHNLIHEDHDMMAFNATVLPDYGX 519  
DB 419 NPIFDVYSTGTTPAAGDNYTIRF--TDNPGWFLHCHIDHLEAGFAVVF-AEDIPD--- 473  
QY 520 NATVEVDPMEEELM 532  
DB 474 --VASANPVPOAM 484

RESULT 6  
US-09-396-260-2  
: Sequence 2, Application US/09396260  
: Patent No. 6184015  
: GENERAL INFORMATION:  
: APPLICANT: Svendsen, Allan  
: APPLICANT: Xu, Feng  
: TITLE OF INVENTION: LACCASE MUTANTS  
: NUMBER OF SEQUENCES: 10  
: CORRESPONDENCE ADDRESS:  
: ADDRESS: No. 61840150 No. 6184015disk of No. 6184015th America, Inc.  
: STREET: 405 Lexington Avenue  
: CITY: New York  
: STATE: NY  
: COUNTRY: USA  
: ZIP: 10174  
: COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FASTSEQ for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/396,260  
: FILING DATE:  
: CLASSIFICATION:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US/09/032,315  
: FILING DATE: 27-FEB-1998  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Rozek, Carol  
: REGISTRATION NUMBER: 36,993  
: REFERENCE/DOCKET NUMBER: 5200,200-US  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 212-867-0123  
: TELEFAX: 212-878-9655  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 499 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-09-396-260-2

Query Match 7.5%; Score 228.5; DB 4; Length 499;  
Best Local Similarity 22.8%; Pred. No. 1,1e-13;  
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

QY 55 IPPVQOPRLVTNPNVNGEIMWEYVEIKPFTHQVYPDLGSADLVGSDGSPGPTFOVPRG 114  
DB 2 IGPVAD--LTITNAA-----VSPDGFSSRAVYVNGGTPGLITGNMG 41  
QY 115 VETVYRFINNAE-----APNSVHLGFSRAAFDGMAE-----DITEPGSF-KDYXX 160  
DB 42 DRFOLNVIDNLNHTMLKSTSIHMGFFOKGT--NMADGPAPINOCPISSGSHFLYDFQV 99  
QY 161 PNRQARFLMTHDHAMHTAENAVRGAGLYMLTPPADALMLPSGYGDFDIPMLITSQ 220  
DB 100 PDQ--AGTFWYHSH--LSTQYCDGLGPFVYVDPNDPADLYVDN---DDTVITLVDM 151  
QY 221 YTANGNLVTTNGELNSFWGDVYIHVNGQ-----PMPFKNVEP-RKYRFRFLDAVSR 270  
DB 152 YHVAAKL-----GPAEPLGADATLNGKGRSPSTTADLSVISVTGKRRYRFLVSLSCP 207  
QY 271 SF-----GLYFADTDALDTRLPKRVIASDSGLLEHPADTSLIYISMAERYEVVDFSD 323  
DB 208 NYTFESIDGHNMTIETDSINTA-----PLVDSIOIFAQORYSFYLE-AN 251  
QY 324 YAGKTIELR--NLG--GSIGIGTDTVDNTDKVMRF--VVAODTQPDTSVVPANLR 375  
DB 222 QAVDNYWIRANPNFGVNGFTGINS-----ALIRYDGAALVEPTTQTSTAPLANEV 303  
QY 376 DVPPSPPTTNTPR-----OFRGRTGPTWTINGVAFADYONRLANVPVGTVER 424  
DB 304 NL-HPLVTTAVGSPVAGVDLAIMMAFNENGTNFINGASTPTPVYLLDIIISGAQNA 362  
QY 425 WELINAGN-----GWTPIHILHVDKVISRTSGNNAKRYMPE 463  
DB 363 QDLPSGSVYSLPSNADIEISFPATAAAGAPHPHILHGAFAVY--RSAGS---TVYND 418  
QY 464 SGL-KDYVWLG---RREYVVEAHYAAPFGVYMFCHNLIHEDHDMMAFNATVLPDYGX 519  
DB 419 NPIFDVYSTGTTPAAGDNYTIRF--TDNPGWFLHCHIDHLEAGFAVVF-AEDIPD--- 473  
QY 520 NATVEVDPMEEELM 532  
DB 474 --VASANPVPOAM 484

RESULT 7  
US-09-576-281-2  
Sequence 2, Application US/09576281  
Patent No. 6277611  
GENERAL INFORMATION:  
APPLICANT: Pedersen, Anders  
APPLICANT: Svendsen, Allan  
APPLICANT: Schneider, Palle  
APPLICANT: Rasmussen, Grethe  
APPLICANT: Cherry, Joel  
TITLE OF INVENTION: LACCASE MUTANTS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 62776110 No. 6277611disk of No. 6277611th America  
STREET: 405 Lexington Avenue  
CITY: New York  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/576,281  
CLASSIFICATION:  
FILING DATE:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/993,318  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Gregg, Valeta A.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 5032.200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 499 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-576-281-2

Query Match 7.5%; Score 228.5; DB 4; Length 499;  
Best local Similarity 22.8%; Pred. No. 1.1e-13;  
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

Db 55 IPVYKQRLVTPVNGQELIWEVEIKPTHOVYEDLGSADLVGDGMSPPGTFVPRG 114  
2 IGVVAD-LTITNAA-----VSPDGFSRQAVVVGSGPGLITGMG 41  
115 VEVVREFINAE-----APNSVHLHGSFSRAFDGAE-----DITEPGSF-KDYY 160  
42 DRQVLAVINDLTHMTKSTIHWGFEQGT--NWDGPAFLNOCPISSGSHFLYDFQY 99  
161 PNRQARTLWYHDHAMHITAENAYRQAGLYMLTDPADALNLPSGYGERDIPMILTSK 220  
100 PQQ--AGTFEYHSH-LSTQYCDGLRGFFVYDNDPADLYVDN-----DDVITLVDM 251  
221 YTANGNLVTTNGELNSFWGVIVHNGO-----PMFKNVEP-RKYRFRFLDAVSR 270  
152 YHVAARL---GAEPFGAATLINSGRSPSTTADLSVSVTPGKRYRFRFLVLSGCDP 207  
271 SF-----GLYFADDAIDATRLPFVYIASDSGLEHPADTSLIYSMARYEVPDFSD 323  
208 NYFESIDGHMMTIETDSITFA-----PLVYDSIQIFAAORYSFVLE-AN 251  
324 YAKTIELR--NLG--GSIGIGITDXYDNTDKVMRF--VVADDTOPDTSVPAANLR 375

Db 252 QAVDNYWIRANPNFNGVGFPGINS-----ALRYDGAAMAEPTTQTSTAPLNEY 303  
376 DVPPSPPTTTPR-----QFRFGRTGPTWTINGVAFADYONRLLANPVGTVER 424  
304 NL-HPLVTTAVPSPVAGVDLAINMAFNFGTNFFINGASFPPVPLLOIISGONA 362  
425 MELINAGN-----GWTPIHHLVDFVYSTSGNNARTVMPYE 463  
363 QDLPSOSVYSLPSNADIEISFPATAAPGAPHFHGHAFAYV-RSAGS--TYNYD 418  
464 SGL-KDVMWLG---RRETVVVEAHYAFPCGYMFHCHNLIHEDHMAAFNATVLPDGY 519  
419 NPIFRODVSTGTPAAGNVITIRF-TDNPGRWFLHCHIDHLEGRFAYV-ADDIP--- 473  
520 NATVEVDPMEEW 532  
474 --VASANPVQAW 484

RESULT 8  
US-08-462-484-2  
Sequence 2, Application US/08462484  
Patent No. 5667531  
GENERAL INFORMATION:  
APPLICANT: Yaver, Debbie Sue  
APPLICANT: Xu, Feng  
APPLICANT: Dalboge, Henrik  
APPLICANT: Schneider, Palle  
APPLICANT: Aaslyng, Dorrit A.  
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 56675310 No. 5667531disk of No. 5667531th America, Inc.  
STREET: 405 Lexington Avenue, Suite 6400  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,484  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/441,147  
FILING DATE: 15-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lowney, Karen A.  
REGISTRATION NUMBER: 31,274  
REFERENCE/DOCKET NUMBER: 4185.010-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 867 0123  
TELEFAX: 212 878 9655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 520 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Polyporus pinsitus  
US-08-462-484-2

Query Match 7.5%; Score 228.5; DB 1; Length 520;  
Best local Similarity 22.8%; Pred. No. 1.2e-13;



ADDRESSEE: Novo Nordisk of North America, Inc.  
STREET: 405 Lexington Avenue, Suite 6400  
COUNTRY: U.S.A.  
ZIP: 10174-6401

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07536  
FILING DATE: 15-June-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/265,534  
FILING DATE: 24-June-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Lowney, Karen A.  
REGISTRATION NUMBER: 31,274  
REFERENCE/DOCKET NUMBER: 4185.204-WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 867 0123  
TELEFAX: 212 878 9655

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 520 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE: Polyporus plinthus  
ORGANISM: Polyporus plinthus  
PCT-US95-07536-2

```

Db 160 ECPKPKGRKRYRFAKAOQYGTSMYHS---HFSAOYG-NGVGAIOINGPA--SLPYDTD 213
OY 207 YGFEFDMILTSKOYANGNLVTTNGELNSFGCDVIHVNGQ-----PWPEKRV---EP 256
Db 214 LGVFPI-----SDYYSSADELVELTRKNSGAFPSDNLFNKTAHPETGCEYANVLTTPG 269
OY 257 RKYRFRFLDAVSRSGLYFAD---TDAIDTRLFPFVYIASDSGLLEHPADTSLYISMAE 313
Db 270 RRRRLRLINTSVENHFOVSLVNHNTMTIIAADVPVNAAMYDS-----LFLGVGQ 318
OY 314 RYEVVEFSDYAK-----TIELRLNGSIGIGTD-----TDYD 348
Db 319 RYDVVLEASTPGNWFNVFTFGGGLCGSRNDYPAIFHYAGAPGPPDEKAPVDHN 378
OY 349 NTD-KVMRFVADDTQOPDTSVVPAULRDVPFSPPTNTPROGRFGRTGPTWTING---- 403
Db 379 CIDLPLMKPVVADVPLSGAKRPDNTLDVTL--DTTGPPLF-----VMKNGSAIN 428
OY 404 -----VAFADYQNRLLANVPVG--TVERWELINANGMT-----HPIH 439
Db 429 IDMGPRVVDYVLTON---TSFPPGYNIVE-----VNGADQMSYWLIENTDPCAPFTLHPMH 481
OY 440 IHLVDEKVIISR-----TSGNNARTVM--PYESGL-----KDYVWLGRRETVVEAHYA 485
Db 482 LHHGHDYVLGRSPDESPPASRERHVPDPARDAGLLSGANPYRVDYTMPLAFGWVVL-AFRA 540
OY 486 PFGVYMFCHNLIHEDHDMMAAFNATVLPDYGNATVFVDPMEL-----WQARPYE 538
Db 541 DNGGAWLFFCHIAMHVSGLGVYLLERADDLRGAVSDADADDLRLCADMRRRYWPTNPYP 600
OY 539 -----LGEFOAGSGQFSVQA 553
Db 601 KSDGGLKRWVEEGEVLVKA 620

RESULT 12
US-09-083-485-2
; Sequence 2, Application US/09083485
; Patent No. 5948121
; GENERAL INFORMATION:
; APPLICANT: Dorrit Aaslyng
; TITLE OF INVENTION: Laccases with Improved Dyeing
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59481210 No. 5948121disk of No. 5948121th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,485
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 4639,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

US-09-083-485-2
Query Match 7.4%; Score 226; DB 2; Length 620;
Best Local Similarity 22.8%; Pred. No. 2,8e-13;
Matches 155; Conservative 80; Mismatches 255; Indels 190; Caps 37;

OY 3 KHTLGAALSLLENSNAVQASPVETSPATGHLFKHVAQISPOYPMFTVPL---PIPPYK 59
Db 2 KSFTISAATLLVGLTSPVAAAP-----PST-----PQDRDLVITREEAAYK 45
OY 60 QPRLVTNPNVGOELWY-----YEVE-----IKPETHOYVPLDGSAD-LVYIDGMS 104
Db 46 AROQSCNTPSN-RACWTIDGYDINTDYEDVSDPGVYVRPYT---LRLTEDMNTGPDGVV 100
OY 105 P-----GPTQVPRGVETVYREFINNAEA-DSNYLHGSFSRAA--PDGMADIT 150
Db 101 KEKVYMLVNNSTIGPTTFADMGDTIQTIVINNLETNGSTIMHGLHOKTGLHDC-ANGIT 159
OY 151 E---PGEFKDYVYPPNRSARTLWYHDHAMHITAEANAQRQAGLYMLTDPADALNLPSG 206
Db 160 ECPKPKGRKRYRFAKAOQYGTSMYHS---HFSAOYG-NGVGAIOINGPA--SLPYDTD 213
OY 207 YGFEFDMILTSKOYANGNLVTTNGELNSFGCDVIHVNGQ-----PWPEKRV---EP 256
Db 214 LGVFPI-----SDYYSSADELVELTRKNSGAFPSDNLFNKTAHPETGCEYANVLTTPG 269
OY 257 RKYRFRFLDAVSRSGLYFAD---TDAIDTRLFPFVYIASDSGLLEHPADTSLYISMAE 313
Db 270 RRRRLRLINTSVENHFOVSLVNHNTMTIIAADVPVNAAMYDS-----LFLGVGQ 318
OY 314 RYEVVEFSDYAK-----TIELRLNGSIGIGTD-----TDYD 348
Db 319 RYDVVLEASTPGNWFNVFTFGGGLCGSRNDYPAIFHYAGAPGPPDEKAPVDHN 378
OY 349 NTD-KVMRFVADDTQOPDTSVVPAULRDVPFSPPTNTPROGRFGRTGPTWTING---- 403
Db 379 CIDLPLMKPVVADVPLSGAKRPDNTLDVTL--DTTGPPLF-----VMKNGSAIN 428
OY 404 -----VAFADYQNRLLANVPVG--TVERWELINANGMT-----HPIH 439
Db 429 IDMGPRVVDYVLTON---TSFPPGYNIVE-----VNGADQMSYWLIENTDPCAPFTLHPMH 481
OY 440 IHLVDEKVIISR-----TSGNNARTVM--PYESGL-----KDYVWLGRRETVVEAHYA 485
Db 482 LHHGHDYVLGRSPDESPPASRERHVPDPARDAGLLSGANPYRVDYTMPLAFGWVVL-AFRA 540
OY 486 PFGVYMFCHNLIHEDHDMMAAFNATVLPDYGNATVFVDPMEL-----WQARPYE 538
Db 541 DNGGAWLFFCHIAMHVSGLGVYLLERADDLRGAVSDADADDLRLCADMRRRYWPTNPYP 600
OY 539 -----LGEFOAGSGQFSVQA 553
Db 601 KSDGGLKRWVEEGEVLVKA 620

RESULT 13
US-08-939-218A-2
; Sequence 2, Application US/08939218A
; Patent No. 5981243
; GENERAL INFORMATION:
; APPLICANT: BERKA, Randy Michael
; APPLICANT: BROWN, Stephen H.
; APPLICANT: XU, Feng
; APPLICANT: SCHNEIDER, Palle
; APPLICANT: OXBEN LL, Karen M.
; APPLICANT: AASLYNG, Dorrit A.
; TITLE OF INVENTION: PURIFIED MYCELIOPTHORA LACCASES AND NUCLEIC
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59812430 No. 5981243disk of No. 5981243th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor

```

CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/939,218A  
FILING DATE: 29-SEPT-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rozek, Carol E.  
REGISTRATION NUMBER: 36,993  
REFERENCE/DOCKET NUMBER: 4184.120-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 620 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-939-218A-2

Query Match 7.48; Score 226; DB 2; Length 620;  
Best Local Similarity 22.8%; Pred. No. 2,8e-13;  
Matches 155; Conservative 80; Mismatches 255; Indels 190; Gaps 37;

QY 3 KHTLGAALSLLENSNAVQSPVETSPATGHLEKRAVAQSPQPMFTVPL---PIRPVK 59  
DB 2 KSTISATLVLGLTSPVAAP-----PST-----PEGRDLVPTTEREAAVK 45  
QY 60 QPRLTVPNGOEWY-----YEVE-----IKPETHOYVPDLGSAD-LVYDGMS 104  
DB 46 ARQSCFTPEN-RACWTDGVDIMTDEVDSDPDGCVARPT---LTLLEVDMNTGPRGVV 100  
QY 105 P-----GPTFOVPRGVETVAFFINNAE-PNSVHLHGSFSA--FDGMAEDIT 150  
DB 101 KERVMVLNNSIIGPTTFADMGDTIQVIVINNLETNGTSIMHGHGKGTMLHOG-ANGIT 159  
QY 151 E-----GSGFDYVYYPNROSARTLMYHDMHMTAENAYRQAGLYMLTDPAEDMLNPSG 206  
DB 160 ECPTRPGGKRVKFAQYGTSTWYS--HFSAYG-NGVGAIDQNGPA--SLPYDITD 213  
QY 207 YGEFDIPMLTSCQYANGMLVTTNGELNSFWGVDIVHNGO-----PWPCKV---EP 256  
DB 214 LGVFP-----SDVYSSADELVELTKNSGAPFSUNVLFNGTAKRPGESEYANVTLP 269  
QY 257 RKYRFRFLDAVSRSEGLYFAD--TDAIDTRLPEKVIASDSGLLEHPADTSLIYSMAE 313  
DB 270 RRHRLRLINTSVENHEQVSLVNHMTIADAMVYVNMATVDS-----LFLGVGQ 318  
QY 314 RYEVVPEFSYAK-----TIELNLGSGTIGIGD-----TDYD 348  
DB 319 KYDVLEASTPGMYWNTFTGGGLCGSRNPYPALIFHYAGAPGPPDEGKAPVDHN 378  
QY 349 NFD-KVMRFVADDTTQPDTSVVPANLADVPPSPNTNTPROGFRGTGPTWING--- 403  
DB 379 CLDLPLNKPVAVARVPLSGAKRDNLDVTL-DTGTGLF-----VMKNGSAIN 428  
QY 404 -----VAFADVONRLANVPV--TVERMELINAGNMT-----HPIH 439  
DB 429 IDMGKPVVDVLTQN--TSFPPGYNIVE-----VNGADQMSYMLIENDPGAPFTLP 481  
QY 440 IHLVDFKVISR-----TSGNNAARTVM--PYESGL-----KQVWVGRETVEA 485  
DB 482 LKHGDFVYVLRSPDESPASNERHYFDPARADGLLSGANPVYRDVTMLPAFGWVVL-AFRA 540

QY 486 PFGVYMFHCHNLIEDHDMMAFNATVLPDYGYNATVFPMEEL-----WQARPYE 538  
DB 541 DNGPAMLFHCHIAHWVSGGVYLERADLRGAVSDADDIDRLCADMRRTWPTNPVP 600  
QY 539 -----LGEFOAQSGQSFVQA 553  
DB 601 KSDGILKRWVEEGEVLVKA 620

## RESULT 14

US-08-462-484-8  
Sequence 8, Application US/08462484  
Patent No. 5667531  
GENERAL INFORMATION:  
APPLICANT: Yaver, Debbie Sue  
APPLICANT: Xu, Feng  
APPLICANT: Dalboge, Henrik  
APPLICANT: Schneider, Paile  
APPLICANT: Aaslyng, Dorrit A.  
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND  
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 56675310 No. 5667531disk of No. 5667531th America, Inc.  
STREET: 405 Lexington Avenue, Suite 6400  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10174-6401

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,484  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/441,147  
FILING DATE: 15-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lowney, Karen A.  
REGISTRATION NUMBER: 31,274  
REFERENCE/DOCKET NUMBER: 4185.010-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 867 0123  
TELEFAX: 212 878 9655  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 511 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-462-484-8

Query Match 7.38; Score 223.5; DB 1; Length 511;  
Best Local Similarity 21.78%; Pred. No. 3,5e-13;  
Matches 117; Conservative 60; Mismatches 227; Indels 135; Gaps 21;

QY 34 HLEKVAQISPOYPMFTVPLPIRPVOPRLTVTPVNGOEWYVEVEIKPETHOYVPDLG 93  
DB 15 HSEGRVS-----AALGPVTD-LTISN-----GIVSPDGF 42  
QY 94 SADLVGDSGSPGPTFOVPRGVETVAFINNAE-----APNSVHLHGSFRAAFDMAE 147  
DB 43 TRAAVLANGVFPGLTGNKDNFOINVIDLNSNETMLKSTSIHMHGFQKGT--NMADG 100  
QY 148 -----DIEPQSGFKRYVYPNROSARTLMYHDMHMTAENAYRQAGLYMLTDPAEDA 200  
DB 101 AAFVNOCPATIGNSFL-YDFTATDQAGTFWYHSH-LSTOYCDGLRGPMVYVDPSPHADL 158

QY 201 LNLPSGYGEFDIPMLITSKOYTANGNLVTNGELNSFWGDVIVHNGQPMPFKKVER-RKY 259  
DB 159 YDV-----DDETTIITLSDMYHTAASL---GAAPFIGSDSTLINGDLAVITVEQCKRY 209  
QY 260 REFRLDAANSRSF-----GLYFADTDAIDTLPKRVIASDSGLLEHPADTSLIYISMA 312  
DB 210 RMLLSLSCDPNVFVSIDGHNMTIIEADAVN-----HEPLTVDSIQIYAG 254  
QY 313 EREYVEF---DESDYAGKTIELRLNGSIGIGTDTDYDNTDKVMRFVVA---DDTTP 365  
DB 255 QRSFVLTAQDIDNFTIRALPSAGTTSFSGGINS-----ALLRYSASEVDPTTTE 306  
QY 366 DTSVVP---ANLRDVPFSPPT-----TNTPROFREGRTGPTWTINGVAFADYONRL 413  
DB 307 TTSVLPLEDEANL---VPLDSPAAPGDPNIGVDYALNLFNFDGTNFEINDVSFVSPTPV 364  
QY 414 LANVPVGTVERMELINAGN-----GWTHPHILVDFEKYSRT 451  
DB 365 LLOILSGTSSAADLPSSGLEAVPSNSTIEISPTITATNAPGAPHPHILGHTEFSVIRTA 424  
QY 452 SGNNARTVMPYESGLKDVWLGRRETVVEAHYAPPGVYMFCHNLIHEDHMMMAFN 510  
DB 425 GSTDTNFVNVR---ADVNTGTIVGDNVTIRFTTDNPGPWFLLCHIDFHLAAGFAIYFS 480

RESULT 15  
US-08-441-147-8  
Sequence 8, Application US/08441147  
Patent No. 5770418  
GENERAL INFORMATION:  
APPLICANT: Taver, Debbie Sue  
APPLICANT: Xu, Feng  
APPLICANT: Dalboge, Henrik  
APPLICANT: Schneider, Palle  
APPLICANT: Aaslyng, Dorrit A.  
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND  
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 57704180 No. 5770418disk of No. 5770418th America, Inc.  
STREET: 405 Lexington Avenue, Suite 6400  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/441,147  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lowmey, Karen A.  
REGISTRATION NUMBER: 31,274  
REFERENCE/DOCKET NUMBER: 4185.010-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 867 0123  
TELEFAX: 212 878 9655  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 511 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-441-147-8

Query Match 7.3%; Score 223.5; DB 1; Length 511;  
Best Local Similarity 21.7%; Pred. No. 3.5e-13;  
Matches 117; Conservative 60; Mismatches 227; Indels 135; Gaps 21;

QY 34 HLEKRYAQISPOXPMETVPLPIPPVKOPRLVTNPVNGQEIWYEVEIKPFTHQVYDGLG 93  
DB 15 HSFGRVS-----AALGPVTD-LTISN-----GDVSPDGR 42  
QY 94 SADLVGIDGSPGPPTQVPRGVETVYREFINNAE-----APNSVHLHGSERAAFDGNAE- 147  
DB 43 TRAAVLANGVFPGLITGNNGDNFOINVDNLSEMTLAKSTSHMHGFFQKGT--NMADG 100  
QY 148 -----DIREPGSFQDYYYPNROSARTLWYHDHAMHITAENAYRQOAGLYMLTDPEDA 200  
DB 101 AAFVNOCPITAGNSFL-YDFTATDQAGTFYHSH-LSTQICDGLRGPMVYVYDPSDPAHL 158  
QY 201 LNLPSGYGEFDIPMLITSKOYTANGNLVTNGELNSFWGDVIVHNGQPMPFKKVER-RKY 259  
DB 159 YDV-----DDETTIITLSDMYHTAASL---GAAPFIGSDSTLINGDLAVITVEQCKRY 209  
QY 260 REFRLDAANSRSF-----GLYFADTDAIDTLPKRVIASDSGLLEHPADTSLIYISMA 312  
DB 210 RMLLSLSCDPNVFVSIDGHNMTIIEADAVN-----HEPLTVDSIQIYAG 254  
QY 313 EREYVEF---DESDYAGKTIELRLNGSIGIGTDTDYDNTDKVMRFVVA---DDTTP 365  
DB 255 QRSFVLTAQDIDNFTIRALPSAGTTSFSGGINS-----ALLRYSASEVDPTTTE 306  
QY 366 DTSVVP---ANLRDVPFSPPT-----TNTPROFREGRTGPTWTINGVAFADYONRL 413  
DB 307 TTSVLPLEDEANL---VPLDSPAAPGDPNIGVDYALNLFNFDGTNFEINDVSFVSPTPV 364  
QY 414 LANVPVGTVERMELINAGN-----GWTHPHILVDFEKYSRT 451  
DB 365 LLOILSGTSSAADLPSSGLEAVPSNSTIEISPTITATNAPGAPHPHILGHTEFSVIRTA 424  
QY 452 SGNNARTVMPYESGLKDVWLGRRETVVEAHYAPPGVYMFCHNLIHEDHMMMAFN 510  
DB 425 GSTDTNFVNVR---ADVNTGTIVGDNVTIRFTTDNPGPWFLLCHIDFHLAAGFAIYFS 480

Search completed: October 3, 2002, 10:54:53  
Job time: 37 sec





**THIS PAGE BLANK (USPTO)**